

**FIGURE 1**

CCAATCGCCCGGTGCGGTGGTGCAGGGTCTCGGGCTAGTC**ATG**GCGTCCCCGTCTCGGAGAC  
TGCAGACTAAACCAGTCATTACTTGTTTCAAGAGCGTTCTGCTAATCTACACTTTTATTTTC  
TGGATCACTGGCGTTATCCTTCTTGACAGTTGGCATTGTTGGGGCAAGGTGAGCCTGGAGAATTA  
CTTTTCTCTTTTAAATGAGAAGGCCACCAATGTCCCCTTCGTGCTCATTGCTACTGGTACCG  
TCATTATTCTTTTGGGCACCTTTGGTTGTTTTGCTACCTGCCGAGCTTCTGCATGGATGCTA  
AAACTGTATGCAATGTTTCTGACTCTCGTTTTTTTGGTCTGAACTGGTCGCTGCCATCGTAGG  
ATTTGTTTTTCAGACATGAGATTAAGAACAGCTTTAAGAATAATTATGAGAAGGCTTTGAAGC  
AGTATAACTCTACAGGAGATTATAGAAGCCATGCAGTAGACAAGATCCAAAATACGTTGCAT  
TGTTGTGGTGTCAACGATTATAGAGATTGGACAGATACTAATTATTACTCAGAAAAAGGATT  
TCCTAAGAGTTGCTGTAACTTGAAGATTGTACTCCACAGAGAGATGCAGACAAAGTAAACA  
ATGAAGGTTGTTTTATAAAGGTGATGACCATTATAGAGTCAGAAATGGGAGTCGTTGCAGGA  
ATTTCCCTTTGGAGTTGCTTGCTTCCAACCTGATTGGAATCTTTCTCGCCTACTGCCWCTCTCG  
TGCCATAACAAATAACCAGTATGAGATAGTG**TAA**CCCAATGTATCTGTGGGCCTATTCCTCT  
CTACCTTTAAGGACATTTAGGGTCCCCCCTGTGAATTAGAAAGTTGCTTGGCTGGAGAACTG  
ACAACACTACTTACTGATAGACCAAAAACTACACCAGTAGGTTGATTCAATCAAGATGTAT  
GTAGACCTAAACTACACCAATAGGCTGATTCAATCAAGATCCGTGCTCGCAGTGGGCTGAT  
TCAATCAAGATGTATGTTTGCTATGTTCTAAGTCCACCTTCTATCCCATTATGTTAGATCG  
TTGAAACCCTGTATCCCTCTGAAACACTGGAAGAGCTAGTAAATTGTAAATGAAGT

**FIGURE 2**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA19902  
><subunit 1 of 1, 245 aa, 1 stop, 1 unknown  
><MW: -1, pI: 8.36, NX(S/T): 1  
MASPSRRLQTKPVITCFKSVLLIYTFIFWITGVILLAVGIWGKVSLENYFSLLNEKATNVPF  
VLIATGTVIILLGTFGCFATCRASAWMLKLYAMFLTLVFLVELVAAIVGFVFRHEIKNSFKN  
NYEKALKQYNSTGDYRSHAVDKIQNTLHCCGVTDYRDWTDNYYSEKGFPSCKLEDCTPQ  
RDADKVNNEGCFIKVMTIIESEMGVVAGISFGVACFQLIGIFLAYCXSRITNNQYEIV

**Important features of the protein:****Signal peptide:**

amino acids 1-42

**Transmembrane domains:**

amino acids 19-42, 61-83, 92-114, 209-230,

**N-glycosylation site.**

amino acids 134-138

**Tyrosine kinase phosphorylation site.**

amino acids 160-168, 160-169

**N-myristoylation site.**

amino acids 75-81, 78-84, 210-216, 214-220, 226-232

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 69-80, 211-222

CCACACGCGTCCGGCGCCGTCGGCCTCGCGTCCATCTTTGCGGTTCTCTCGGACCTGTCAACAA  
GGAGTCGCGCCGCCGCCGCCGCCCTCCCTCCGGTGGGCCCGGGAGGTAGAGAAAGTCAGT  
GCCACAGCCCGACCGCGCTGCTCTGAGCCCTGGGCACGCGGAACGGGAGGGAGTCTGAGGGT  
TGGGGACGTCTGTGAGGGAGGGGAACAGCCGCTCGAGCCTGGGGCGGGCGGACCGGACTGGC  
GCCGGGGTAGGCTCTGGAAAGGGCCCCGGGAGAGAGGTGGCGTTGGTCAGAACCTGAGAAACA  
GCCGAGAGGTTTTTCCACCGAGGCCCCGCGCTTGAGGGATCTGAAGAGGTTCCCTAGAAGAGGGT  
GTTCCCTCTTTTGGGGGTCTCACCAGAAGAGGTTCTTGGGGGTGCGCCCTTCTGAGGAGGCT  
GCGGCTAACAGGGCCCAGAACTGCCATTGGATGTCCAGAATCCCCTGTAGTTGATAATGTTG  
GGAATAAGCTCTGCAACTTTCTTTGGCATTCACTTGTAAAAACAAATAGGATGCAAATTCC  
TCAACTCCAGGTTATGAAAACAGTACTTGAAAACTGAAAACTACCTAAATGATCGTCTTTG  
GTTGGGCCGTGTTCTTAGCGAGCAGAAGCCTTGGCCAGGGTCTGTTGTTGACTCTCGAAGAG  
CACATAGCCCACTTCTAGGGACTGGAGGTGCCGCTACTACCATGGGTAATTCCTGTATCTG  
CCGAGATGACAGTGGAACAGATGACAGTGTTGACACCCAACAGCAACAGGCCGAGAACAGTG  
CAGTACCCACTGCTGACACAAGGAGCCAACCACGGGACCCTGTTGCGGCCACCAAGGAGGGGC  
CGAGGACCTCATGAGCCAAGGAGAAAGAAACAAAATGTGGATGGGCTAGTGTTGGACACACT  
GGCAGTAATACGGACTCTTGATAGATAAGTAAGTATCTGACTCACGGTCACCTCCAGTGGAAT  
GAAAAGTGTTCTGCCCGBAACCATGACTTTAGGACTCCTTCAGTTCCTTTAGGACATACTCG  
CCAAGCCTTGCTGCTCACAGGGCAAAGGAGAATATTTTAATGCTCCGCTGATGGCAGAGTAA  
TGATAAGATTTGATGTTTTTGCTTGCTGTCATCTACTTTGTCTGGAAATGTCTAAATGTTTC  
TGTAGCAGAAAACACGATAAAGCTATGATCTTTATTAGAG

## **FIGURE 4**

MIVFGWAVFLASRSLGQGLLLTLEEHIAHFLGTGGAATTMGNSCICRDDSGTDDSVDTQQQQ  
AENSAVPTADTRSQPRDPVRPPRRGRGPHEPRRKKQNV DGLVLDTLAVIRTLVDKO

**Signal peptide:**

amino acids 1-16

**Casein kinase II phosphorylation site.**

amino acids 22-26, 50-54, 113-117

**N-myristoylation site.**

amino acids 18-24, 32-38, 34-40, 35-41, 51-57

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



**FIGURE 5**

GGCACGAGGCGCTGTCCACCCGGGGGCGTGGGAGTGAGGTACCAGATTAGCCCATTTGGCC  
CCGACGCCTCTGTTCTCGGAATCCGGGTGCTGCGGATTGAGGTCCCGGTTCCTAACGGACTG  
CAAGATGGAGGAAGGCGGGAACCTAGGAGGCCTGATTAAGATGGTCCATCTACTGGTCTTGT  
CAGGTGCCTGGGGCATGCAAATGTGGGTGACCTTCGTCTCAGGCTTCCTGCTTTTCCGAAGC  
CTTCCCCGACATACCTTCGGACTAGTGCAGAGCAAACCTCTTCCCCTTCTACTTCCACATCTC  
CATGGGCTGTGCCTTCATCAACCTCTGCATCTTGGCTTCACAGCATGCTTGGGCTCAGCTCA  
CATTCTGGGAGGCCAGCCAGCTTTACCTGCTGTTCTGAGCCTTACGCTGGCCACTGTCAAC  
GCCCCTGGCTGGAACCCCGCACACAGCTGCCATGTGGGGCCCTGCAAACCGTGGAGAAGGA  
GCGAGGCCTGGGTGGGGAGGTACCAGGCAGCCACCAGGGTCCCGATCCCTACCGCCAGCTGC  
GAGAGAAGGACCCCAAGTACAGTGCTCTCCGCCAGAATTTCTTCCGCTACCATGGGCTGTCC  
TCTCTTTGCAATCTGGGCTGCGTCCTGAGCAATGGGCTCTGTCTCGCTGGCCTTGCCCTGGA  
AATAAGGAGCCTCTAGCATGGGCCCTGCATGCTAATAAATGCTTCTTCAGAAATGAAAAAA  
AAAAAAAAAAAA

**FIGURE 6**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56107

<subunit 1 of 1, 231 aa, 1 stop

<NX(S/T): 0

MEEGGNLGGLIKMVHLLVLSGAWGMQMWVTFVSGFLLFRSLPRHTFGLVQSKLFPFYFHISM  
GCAFINLCILASQHAWAQLTFWEASQLYLLFLSLTLATVNARWLEPRTTAAMWALQTVEKER  
GLGGEVPGSHQGPDPYRQLREKDPKYSALRQNFFRYHGLSSLCNLGCVLSNGLCLAGLALEIRSL

**Signal peptide:**

amino acids 1-24

**Transmembrane domain:**

amino acids 86-103, 60-75

**Casein kinase II phosphorylation site.**

amino acids 82-86

**Tyrosine kinase phosphorylation site.**

amino acids 144-151

**N-myristoylation site.**

amino acids 4-10, 5-11, 47-53, 170-176, 176-182

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 54-65

**G-protein coupled receptors proteins.**

amino acids 44-85

## FIGURE 7

AATTTCAGATTTTTAAAGCCCATTTCTGCAGTGGAAATTTTCATGAAC TAGCAAGAGGACACCATCTT  
CTTGTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGGGAAAAATGCTCTTTTGGGT  
GCTAGGCCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAAC TAAAGATTGAAG  
ACATCACTGATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAAACTTGGCAGCC  
AGAACTTTTGATAAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAAC  
AGCTTTAAAGGCAGAAACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAG  
AGAATGTCAAGAGGACTGCCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGT  
CTGATCAATAATGCTGGTGTTCCTCGCGTGTCTGGCTCCCACTGACTGGCTGACACTAGAGGA  
CTACAGAGAACCTATTGAAGTGAACCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTC  
CTTTGGTCAAGAAAGCTCAAGGGAGAGTTATTAATGTCTCCAGTGTGGAGGTCGCCTTGCA  
ATCGTTGGAGGGGGCTATACTCCATCCAAATATGCAGTGGAAAGGTTTCAATGACAGCTTAAG  
ACGGGACATGAAAGCTTTTGGTGTGCACGTCTCATGCATTGAACCAGGATTGTTCAAAACAA  
ACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAACTCGCCATTTGGGAGCAGCTGTCTCCA  
GACATCAAACAACAATATGGAGAAGGTTACATTGAAAAAAGTCTAGACAAACTGAAAGGCAA  
TAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTGCATGGACCACGCTCTAACAA  
GTCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAATTTTCTGGATACCTCTG  
TCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAAGCAGAGCTGGCTAA  
TCCCAAGGCAGTG TGA CTCACTAGCTAACCACAAATGTCTCCTCCAGGCTATGAAATTGGCCGAT  
TTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACTCATTTAGA  
TCGTGCTTATTTGGATTGCAAAGGGAGTCCCACCATCGCTGGTGGTATCCAGGGTCCCTG  
CTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCTGT  
ATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATG  
ATCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCCAGCATTTACAGTAACTTGTGAATGTT  
AAGTATCATCTCTTATCTAAATATTTAAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAA

**FIGURE 8**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56406

><subunit 1 of 1, 319 aa, 1 stop

><MW: 35227, pI: 8.97, NX(S/T): 3

MLFWVLGLLILCGFLWTRKGLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLT  
ESGSTALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDW  
LTLEDYREPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGF  
NDSLRRDMKAFGVHVSCIEPGLFECTNLADPVKVIEKKLAIWEQLSPDIKQQYGEGYIEKSLD  
KLKGNKSYVNMDLSPVVECMDHALTSLEPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQK  
AELANPKAV

**Important features of the protein:****Signal peptide:**

amino acids 1-17

**Transmembrane domain:**

amino acids 136-152

**N-glycosylation sites.**

amino acids 161-163, 187-190 and 253-256

**Glycosaminoglycan attachment site.**

amino acids 39-42

**N-myristoylation sites.**

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

**FIGURE 9**

GCGGGCTGTTGACGGCGCTGCG**ATG**GCTGCCTGCGAGGGCAGGAGAAGCGGAGCTCTCGGTT  
 CCTCTCAGTCGGACTTCCTGACGCCGCCAGTGGGCGGGGCCCTTGGGCGGTCGCCACCACT  
 GTAGTCATGTACCCACCGCCGCCGCCGCCCTCATCGGGACTTCATCTCGGTGACGCTGAG  
 CTTTGGCGAGAGCTATGACAACAGCAAGAGTTGGCGGGCGGCTCGTGCTGGAGGAAATGGA  
 AGCAACTGTCGAGATTGCAGCGGAATATGATTCTCTTCTCCTTGCCTTTCTGCTTTTCTGT  
 GGACTCCTCTTCTACATCAACTTGGCTGACCATTGGAAGCTCTGGCTTTCAGGCTAGAGGA  
 AGAGCAGAAGATGAGGCCAGAAATTGCTGGGTAAACCAGCAAATCCACCCGTCTTACCAG  
 CTCCTCAGAAGGCGGACACCGACCCTGAGAACTTACCTGAGATTTCTGTACAGAAAGACACAA  
 AGACACATCCAGCGGGGACCACCTCACCTGCAGATTAGACCCCCAAGCCAAGACCTGAAGGA  
 TGGGACCCAGGAGGAGGCCACAAAAGGCAAGAAGCCCCCTGTGGATCCCCGCCCCGGAAGGAG  
 ATCCGCAGAGGACAGTCATCAGCTGGAGGGGAGCGGTGATCGAGCCTGAGCAGGGCACCAGG  
 CTCCCTTCAAGAAGAGCAGAAGTGCCACCAAGCCTCCCTGCCACCGGCCAGGACACAGGG  
 CACACCAGTGCATCTGAACTATCGCCAGAAGGGCGTGATTGACGTCTTCTGTCATGCATGGA  
 AAGGATACCGCAAGTTTGCATGGGGCCATGACGAGCTGAAGCCTGTGTCCAGGTCTTTCAGT  
 GAGTGGTTTGGCCTCGGTCTCACACTGATCGACGCGCTGGACACCATGTGGATCTTGGGTCT  
 GAGGAAAGAATTTGAGGAAGCCAGGAAGTGGGTGTGCAAGAAGTTACACTTTGAAAAGGACG  
 TGGACGTCAACCTGTTTGAGAGCACGATCCGCATCCTGGGGGGGCTCCTGAGTGCCTACCAC  
 CTGTCTGGGGACAGCCTCTTCTGAGGAAAGCTGAGGATTTTGGAAATCGGCTAATGCCTGC  
 CTTCAGAACACCATCCAAGATTCTTACTCGGATGTGAACATCGGTACTGGAGTTGCCACC  
 CGCCACGGTGGACCTCCGACAGCACTGTGGCCGAGGTGACCAGCATTACAGCTGGAGTTCGG  
 GAGCTCTCCCGTCTCACAGGGGATAAGAAGTTTCAGGAGGCAAGTGGAGAAGGTGACACAGCA  
 CATCCACGGCCTGTCTGGGAAGAAGGATGGGCTGGTGCCCATGTTTCATCAATACCCACAGTG  
 GCCTCTTACCCACCTGGGCGTATTACGCTGGGCGCCAGGGCCGACAGCTACTATGAGTAC  
 CTGCTGAAGCAGTGGATCCAGGGCGGGAAGCAGGAGACACAGCTGCTGGAAGACTACGTGGA  
 AGCCATCGAGGGTGTGAGAACGCACCTGCTGCGGCACTCCGAGCCAGTAAGCTCACCTTTG  
 TGGGGGAGCTTGCCACGGCCGCTTCAAGTGCACAGATGGACCACCTGGTGTGCTTCTGCCA  
 GGGACGCTGGCTCTGGGCGTCTACCACGGCCTGCCCGCCAGCCACATGGAGCTGGCCAGGA  
 GCTCATGGAGACTTGTACCAGATGAACCGGCAGATGGAGACGGGGCTGAGTCCCGAGATCG  
 TGCATTTCAACCTTTACCCCCAGCCGGGCGCTCGGGACGTGGAGGTCAAGCCAGCAGACAGG  
 CACAACCTGCTGCGGCCAGAGACCGTGGAGAGCCTGTTCTACCTGTACCGCGTCACAGGGGA  
 CCGCAAATACCAGGACTGGGGCTGGGAGATTCTGCAGAGCTTCAGCCGATTACACGGGTCC  
 CCTCGGGTGGCTATTCTTCCATCAACAATGTCCAGGATCCTCAGAAGCCCGAGCCTAGGGAC  
 AAGATGGAGAGCTTCTTCTGAGGAGACGCTCAAGTATCTGTTCTTGTCTTCTCCGATGA  
 CCCAAACCTGCTCAGCCTGGACGCCTACGTGTTCAACACCGAAGCCCACCCTCTGCCTATCT  
 GGACCCCTGCC**TAG**GGTGGATGGCTGCTGGTGTGGGGACTTCGGGTGGGCAGAGGCACCTTG  
 CTGGGTCTGTGGCATTTCCTCAAGGGCCACGTAGCACCGGCAACCGCCAAGTGGCCCAGGCT  
 CTGAACTGGCTCTGGGCTCCTCCTCGTCTCTGCTTTAATCAGGACACCGTGAGGACAAGTGA  
 GGCCGTGAGTCTTGGTGTGATGCGGGGTGGGCTGGGCGCTGGAGCCTCCGCCTGCTTCTC  
 CAGAAGACACGAATCATGACTCACGATTGCTGAAGCCTGAGCAGGTCTCTGTGGGCCGACCA  
 GAGGGGGGCTTCGAGGTGGTCCCTGGTACTGGGGTGACCGAGTGGACAGCCCAGGGTGCAGC  
 TCTGCCCCGGGCTCGTGAAGCCTCAGATGTCCCAATCCAAGGGTCTGGAGGGGCTGCCGTGA  
 CTCCAGAGGCCTGAGGCTCCAGGGCTGGCTCTGGTGTTCACAAGCTGGACTCAGGGATCCTC  
 CTGGCCGCCCCGACAGGGGCTTGAGGGCTGGACGGCAAGTCCGTCTAGCTCACGGGCCCCCT  
 CCAGTGGAATGGGTCTTTTCGGTGGAGATAAAAGTTGATTTGCTCTAACCGCAA

**FIGURE 10**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56529

><subunit 1 of 1, 699 aa, 1 stop

><MW: 79553, pI: 7.83, NX(S/T): 0

MAACEGRRSGALGSSQSDFLTTPFVGAPWAVATTVMYPPPPPPPHRDFISVTLSEGESYDN  
SKSWRRRSCWRKWKQLSRLQRMILFLLAFLLFCGLLFYINLADHWKALAFRLEEEQKMRPE  
IAGLKPANPPVLPAPQKADTDPENLPEISSQKTQRHIQRGPPHLQIRPPSQDLKDGTQEEAT  
KRQEAPVDPRPEGDPQRTVISWRGAVIEPEQGTELPSRRAEVPTKPPLPPARTQGTPVHLNY  
RQKGVIDVFLHAWKGYRKFAWGHDELKPVSRSFSEWFGLGLTLIDALDTMWILGLRKEFEEA  
RKWVSKKLHFEKDVDVNLFESTITIRILGGLLSAYHLSGDSLFLRKAEDFGNRLMPAFRTPSKI  
PYSDVNIGTGVAHPPRWTSDSTVAEVTISIQLFREL SRLTGDKKFQEAVEKVTQHIHGLSGK  
KDGLVPMFINTHSGLFTHLGVFTLGARADSYEYLLKQWIQGGKQETQLLEDYVEAIEGVRT  
HLLRHSEPSKLT FVGELAHGRFSAKMDHLVCFLPGTLALGVYHGLPASHMELAQELMETCYQ  
MNRQMETGLSPEIVHFNLYPQPGRRDVEVKPADRHNLRLPETVESLFYLYRVTGDRKYQDWG  
WEILQSFSRFRTRVPSGGYSSINNVQDPQKPEPRDKMESFFLGETLKYLFLLFSDDPNLLSLD  
AYVFNTEAHPLPIWTPA

**Important features of the protein:**

**Transmembrane domain:**

amino acids 21-40 and 84-105 (type II)

## FIGURE 11

GGCGCCGCGTAGGCCGCGGAGGCCCGGGCGGGCTGCGAGCGCCTGCCCATGCGCCGC  
CGCCTCTCCGCACGATGTTCCCCTCGCGGAGGAAAGCGGCGCAGCTGCCCTGGGAGGACGGC  
AGGTCCGGGTTGCTCTCCGGCGGCCTCCCTCGGAAGTGTTCCGTCTTCCACCTGTTTCGTGGC  
CTGCCTCTCGCTGGGCTTCTTCTCCCTACTCTGGCTGCAGCTCAGCTGCTCTGGGGACGTGG  
CCCGGGCAGTCAGGGGACAAGGGCAGGAGACCTCGGGCCCTCCCCGTGCCTGCCCCCAGAG  
CCGCCCCCTGAGCACTGGGAAGAAGACGCATCCTGGGGCCCCCACC GCCTGGCAGTGCTGGT  
GCCCTTCCGCGAACGCTTCGAGGAGCTCCTGGTCTTCGTGCCCCACATGCGCCGCTTCTTGA  
GCAGGAAGAAGATCCGGCACCACATCTACGTGCTCAACCAGGTGGACCACTTCAGGTTCAAC  
CGGGCAGCGCTCATCAACGTGGGCTTCTGGAGAGCAGCAACAGCACGGACTACATTGCCAT  
GCACGACGTTGACCTGCTCCCTCTCAACGAGGAGCTGGACTATGGCTTTCCTGAGGCTGGGC  
CCTTCCACGTGGCCTCCCCGGAGCTCCACCCTCTCTACCACTACAAGACCTATGTCGGCGGC  
ATCCTGCTGCTCTCCAAGCAGCACTACCGGCTGTGCAATGGGATGTCCAACCGCTTCTGGGG  
CTGGGGCCGCGAGGACGACGAGTTCTACCGGCGCATTAAAGGGAGCTGGGCTCCAGCTTTTCC  
GCCCCCTCGGGAATCACAACCTGGGTACAAGACATTTGCCCACCTGCATGACCCAGCCTGGCGG  
AAGAGGGACCAGAAGCGCATCGCAGCTCAAAAACAGGAGCAGTTCAAGGTGGACAGGGAGGG  
AGGCCTGAACACTGTGAAGTACCATGTGGCTTCCCGCACTGCCCTGTCTGTGGGCGGGGCC  
CCTGCACTGTCCTCAACATCATGTTGGACTGTGACAAGACCGCCACACCCTGGTGCACATTC  
AGCTGAGCTGGATGGACAGTGAGGAAGCCTGTACCTACAGGCCATATTGCTCAGGCTCAGGA  
CAAGGCCTCAGGTCGTGGGCCCAGCTCTGACAGGATGTGGAGTGGCCAGGACCAAGACAGCA  
AGCTACGCAATTGCAGCCACCCGGCCGCCAAGGCAGGCTTGGGCTGGGCCAGGACACGTGGG  
GTGCCTGGGACGCTGCTTGCCATGCACAGTGATCAGAGAGAGGCTGGGGTGTGTCCTGTCCG  
GGACCCCCCTGCCTTCTGCTCACCTACTCTGACCTCCTTCACGTGCCCAGGCCTGTGGG  
TAGTGGGGAGGGCTGAACAGGACAACCTCTCATCACCTACTCTGACCTCCTTCACGTGCCC  
AGGCCTGTGGGTAGTGGGGAGGGCTGAACAGGACAACCTCTCATCACCCCCAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 12**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56531

><subunit 1 of 1, 327 aa, 1 stop

><MW: 37406, pI: 9.30, NX(S/T): 1

MFPSRRKAAQLPWEDGRSGLLSGGLPRKCSVFHLEFVACLSLGGFSLWLQLSCSGDVARAVR  
GQQQETSGPPRACPPEPPPEHWEEDASWGPHRLAVLVFPRERFEELLVFVPHMRRFLSRKKI  
RHHIYVLNQVDHFRFNRAALINVGFLSSNSTDYIAMHDVDLLPLNEELDYGFPEAGPFHVA  
SPELHPLYHYKTYVGGILLLSKQHYRLCNGMSNRFWGWGREDDFYRRIKGAGLQLFRPSGI  
TTGYKTFRHLHDPARWRKRDQKRIAAQKQEQFKVDREGGLNTVKYHVASRTALSVGGAPCTVL  
NIMLDCDKTATPWCTFS

**Signal peptide:**

amino acids 1-42

**Transmembrane domain:**

amino acids 29-49 (type II)

**N-glycosylation site.**

amino acids 154-158

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 27-31

**Tyrosine kinase phosphorylation site.**

amino acids 226-233

**N-myristoylation site.**

amino acids 19-25, 65-71, 247-253, 285-291, 303-309, 304-310



**FIGURE 13**

CAATGTTTGCCTATCCACCTCCCCCAAGCCCCTTTACCT**ATG**CTGCTGCTAACGCTGCTGCT  
GCTGCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCAGAAAGTCT  
CTTCTGCCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTC  
CTGCCTCATCGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGT  
GGGGGATGGC**TAA**GAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGG  
GCTGCCTTATTTAAAGTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGC  
CCTGTTCAATTAAGAAATTGTTCCCTTCCCCTGTGTTCAATGTTTGTAAAGATTGTTCTGTGT  
AAATATGTCTTTATAATAAACAGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 14**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56862

<subunit 1 of 1, 73 aa, 1 stop

<MW: 7879, pI: 7.21, NX(S/T): 0

MLLLTLLLLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFPPSFLCLLPHRPAMTCSQAQ

PRGEGEKVGDG

**Important features:**

**Signal peptide:**

amino acids 1-15

**Growth factor and cytokines receptors family:**

amino acids 3-18

14/249

### FIGURE 15

GGGACCCATGCGGCCGTGACACCCCGGCTCCCTAGAGGCCACGCGCAGCCGACGAGGACAAAG  
GAGCATGTCCGCGCCGGGGAAGGCCCGTCTCTCCGGCCGCCATAAAGCTCCGGTCCGCCGCTGG  
GCCCCGCGCCGCTCTCTGCCCGCCGGGCTCCGGGGCGGCCCGCTAGGCCAGTGCGCCGCCG  
CTCGCCCCGACAGGCCCCGGCCCGCAGC**ATG**GAGCCACCCGGACGCCGGCGGGGGCCGCGCGCA  
GCCGCCGCTGTTGCTGCCGCTCTCGCTGTTAGCGCTGCTCGCGCTGCTGGGAGGCGGCGGCG  
GCGGCGGCGCCGCGGCGCTGCCCGCCGGCTGCAAGCACGATGGGCGGCCCCGAGGGGCTGGC  
AGGGCGGCGGGCGCCGCCGAGGGCAAGGTGGTGTGCAGCAGCCTGGAACCTCGCGCAGGTCTT  
GCCCCCAGATACTCTGCCCAACCGCACGGTCACCCTGATTCTGAGTAACAATAAGATATCCG  
AGCTGAAGAATGGCTCATTTTTCTGGGTAAAGTCTCCTTGAAAGATTGGACCTCCGAAACAAT  
CTTATTAGTAGTATAGATCCAGGTGCCTTCTGGGGACTGTCATCTCTAAAAAGATTGGATCT  
GACAAACAATCGAATAGGATGTCTGAATGCAGACATATTTTCGAGGACTCACCAATCTGGTTC  
GGCTAAACCTTTTCGGGGAATTTGTTTTCTTCATTATCTCAAGGAACTTTTGATTATCTTGCG  
TCATTACGGTCTTTGGAATTCCAGACTGAGTATCTTTTGTGTGACTGTAACATACTGTGGAT  
GCATCGCTGGGTAAAGGAGAAGAACATCACGGTACGGGATACCAGGTGTGTTTATCCTAAGT  
CACTGCAGGCCCAACCAGTCACAGGCGTGAAGCAGGAGCTGTTGACATGCGACCCTCCGCTT  
GAATTGCCGTCTTTCTACATGACTCCATCTCATCGCCAAGTTGTGTTTGAAGGAGACAGCCT  
TCCTTTCCAGTGCATGGCTTCATATATTGATCAGGACATGCAAGTGTGTGGTATCAGGATG  
GGAGAATAGTTGAAACCGATGAATCGCAAGGTATTTTTGTTGAAAAGAACATGATTCACAAC  
TGCTCCTTGATTGCAAGTGCCCTAACCATTTCTAATATTCAGGCTGGATCTACTGGAATTG  
GGGCTGTCATGTCCAGACCAACGTGGGAATAAATACGAGGACTGTGGATATTGTGGTATTAG  
AGAGTTCTGCACAGTACTGTCTCCAGAGAGGGTGGTAAACAACAAAGGTGACTTCAGATGG  
CCCAGAACATTGGCAGGCATTACTGCATATCTGCAGTGTACGCGGAACACCCATGGCAGTGG  
GATATATCCCGGAAACCCACAGGATGAGAGAAAAGCTTGGCGCAGATGTGATAGAGGTGGCT  
TTTGGGCAGATGATGATTATTCTCGCTGTCAGTATGCAAATGATGTCACTAGAGTTCTTTAT  
ATGTTTAATCAGATGCCCCCTCAATCTTACCAATGCCGTGGCAACAGCTCGACAGTTACTGGC  
TTACACTGTGGAAGCAGCCAACTTTTCTGACAAAATGGATGTTATATTTGTGGCAGAAATGA  
TTGAAAAATTTGGAAGATTTACCAAGGAGGAAAAATCAAAAGAGCTAGGTGACGTGATGGTT  
GACATTGCAAGTAACATCATGTTGGCTGATGAACGTGTCTGTGGCTGGCGCAGAGGGAAGC  
TAAAGCCTGCAGTAGGATTGTGCAGTGTCTTCAGCGCATTGCTACCTACCGGCTAGCCGGTG  
GAGCTCACGTTTATTCAACATATTCACCCAATATTGCTCTGGAAGCTTATGTATCAAGTCT  
ACTGGCTTCACGGGGATGACCTGTACCGTGTTCCAGAAAGTGGCAGCCTCTGATCGTACAGG  
ACTTTCGGATTATGGGAGGCGGGATCCAGAGGGAAACCTGGATAAGCAGCTGAGCTTTAAGT  
GCAATGTTTCAAATACATTTTCGAGTCTGGCACTAAAGGTATGTTACATTCTGCAATCATTT  
AAGACTATTTACAGT**TAA**ATTAGAATGCTCCAAATGTTCTGCTTCGCAAAATAACCTTATTA  
AAAGATTTTTTTTTTGCAAGGAAGATAGGTATTATTGCTTTTGCTACTGTTTTAAAGAAAATA  
ACCAGGAAGAACTGCATTACGACTTTCAAGGGCCCTAGGCATTTTTGCCTTTGATTCCTTTT  
CTTCACATAAAAATATCAGAAATTACATTTTATAACTGCAGTGGTATAAATGCAAATATACT  
ATTGTTACATAGTGAAAAAATTTTATTGACTTAAAAGTTTATTTATTTGTTTTTTTTGCTCCT  
GATTTTAAGACATAAAGATGTTTTTCATGGGCCCTTAAAGTATCATGAGCCTTTGGCACTGC  
GCCTGCCAAGCCTAGTGGAGAAAGTCAACCTTGAGACCAGGTGTTTAATCAAGCAAGCTGTAT  
ATCAAAATTTTTTGGCAGAAAAACACAAATATGTCATATATCTTTTTTTTAAAAAAAGTATTTCA  
TTGAAGCAAGCAAAATGAAAGCATTTTTTACTGATTTTTTAAATTTGGTGCTTTAGATATATTT  
GACTACACTGTATTGAAGCAAATAGAGGAGGCACAACCTCCAGCACCTAATGGAACCACATT  
TTTTTCACTTAGCTTTCTGTGGGCATGTGTAATTGTATTCTCTGCGGTTTTTAATCTCACAG  
TACTTTATTTCTGTCTTGTCCTCAATAATATCACAAACAATATTCCAGTCATTTTAATGGC  
TGCATAATAACTGATCCAACAGGTGTTAGGTGTTCTGGTTTTAGTGTGAGCACTCAATAAATA  
TTGAATGAATGAACGAAAAAAAAAAAAAAAAA

**FIGURE 16**

MEPPGRRRRGRAQPPLLLPLSLLALLALLGGGGGGGAAALPAGCKHDGRPRGAGRAAGAAEGK  
VVCSSLELAQVLPDPTLPNRTVTILISNNKISELKNGSFSGLSLLERLDLRNNLISSIDPGA  
FWGLSSSLKRLDLTNNRIGCLNADIFRGLTNLVRLNLSGNLFSSLSQGTFDYLASLSLEFQT  
EYLLCDCNILWMHRWVKEKNITVRDTRCVYPKSLQAQPVTGVKQELLTCDPPELPSFYMTP  
SHRQVVFEGLSLPFQCMASYIDQDMQVLWYQDGRIVETDESQGI FVEKNMIHNCSLIASALT  
ISNIQAGSTGNWGCHVQTKRGNNTRTVDIVVLESSAQYCPPERVVNNKGDFRWPRTLAGITA  
YLQCTRNTHGSGIYPGNPQDERKAWRRCDRGGFWADDDYSRCQYANDVTRVLYMFNQMPNLN  
TNAVATARQLLAYTVEAANFSDKMDVIFVAEMIEKFGRFTKEEKSKELGDMVDIASNIMLA  
DERVLWLAQREAKACSRIVQCLQRIATYRLAGGAHVYSTYSPNIALEAYVIKSTGFTGMTCT  
VFQKVAASDRTGLSDYGRDPEGNLDKQLSFKCNVSNTFSSLALKVCYILQSFKTIYS

**Signal peptide:**

amino acids 1-33

**Transmembrane domain:**

amino acids 13-40 (type II)

**N-glycosylation site.**

amino acids 81-85, 98-102, 159-163, 206-210, 301-305, 332-336,  
433-437, 453-457, 592-596

**N-myristoylation site.**

amino acids 29-35, 30-36, 31-37, 32-38, 33-39, 34-40, 51-57,  
57-63, 99-105, 123-129, 142-148, 162-168, 317-323, 320-326,  
384-390, 403-409, 554-560

**FIGURE 17**

GCGTGGGG**ATG**TCTAGGAGCTCGAAGGTGGTGTCTGGGCCTCTCGGTGCTGCTGACGGCGGCC  
ACAGTGGCCGGCGTACATGTGAAGCAGCAGTGGGACCAGCAGAGGCTTCGTGACGGAGTTAT  
CAGAGACATTGAGAGGCAAATTCGGAAAAAAGAAAACATTCGTCTTTTGGGAGAACAGATTA  
TTTTGACTGAGCAACTTGAAGCAGAAAGAGAGAAGATGTTATTGGCAAAGGATCTCAAAAA  
TCAT**TGA**CTTGAATGTGAAATATCTGTTGGACAGACAACACGAGTTTGTGTGTGTGTGTTGAT  
GGAGAGTAGCTTAGTAGTATCTTCATCTTTTTTTTTTGGTCACTGTCCTTTTAAACTTGATCA  
AATAAAGGACAGTGGGTCAATAAAGTTACTGCTTTCAGGGTCCCTTATATCTGAATAAAGGA  
GTGTGGGCAGACACTTTTTGGAAGAGTCTGTCTGGGTGATCCTGGTAGAAGCCCCATTAGGG  
TCACTGTCCAGTGCTTAGGGTTGTTACTGAGAAGCACTGCCGAGCTTGTGAGAAGGAAGGGA  
TGGATAGTAGCATCCACCTGAGTAGTCTGATCAGTCGGCATGATGACGAAGCCACGAGAACA  
TCGACCTCAGAAGGACTGGAGGAAGGTGAAGTGAGGGAGAGACGCTCCTGATCGTCGAATCC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## **FIGURE 18**

MSRSSKVVLGLSVLLTAATVAGVHVKQQWDQQRLRDGVIRDIERQIRKKENIRLLGEQIILT  
EQLEAEREKMLLAKGSQKS

**Signal peptide:**

amino acids 1-21

18/249

## FIGURE 19

CTGTCTGCTTTTGCTTTCAGCCGACAGTCGCCACTGGCTGACCTGAGGTGCTCTTACAGCCTGTTC  
CAAGTGTGGCTTAATCCGTCTCCACCACCAGATCTTTCTCCGTGGATTCCCTCTGCTAAGACC  
GCTGCCATGCCAGTGACGGTAACCCGCACCACCATCACAACCACCACGACGTCATCTTCGGG  
CCTGGGGTCCCCCATGATCGTGGGGTCCCCTCGGGCCCTGACACAGCCCCTGGGTCTCCTTCGC  
CTGCTGCAGCTGGTGTCTACCTGCGTGGCCTTCTCGCTGGTGGCTAGCGTGGGCGCCTGGAC  
GGGGTCCATGGGCAACTGGTCCATGTTACCTGGTGTCTTCTGCTTCTCCGTGACCTGATCA  
TCCTCATCGTGGAGCTGTGCGGGCTCCAGGCCCGCTTCCCCCTGTCTTGGCGCAACTTCCCC  
ATCACCTTCGCCTGCTATGCGGCCCTCTTCTGCCTCTCGGCCTCCATCATCTACCCACCAC  
CTATGTCCAGTTCCTGTCCACGGCCGTTTCGCGGGACCACGCCATCGCCGCCACCTTCTTCT  
CCTGCATCGCGTGTGTGGCTTACGCCACCGAAGTGGCCTGGACCCGGGCCCCGGCCCCGGCGAG  
ATCACTGGCTATATGGCCACCGTACCCGGGGCTGCTGAAGGTGCTGGAGACCTTCGTTGCCTG  
CATCATCTTCGCGTTCATCAGCGACCCCAACCTGTACCAGCACCAGCCGGCCCTGGAGTGGT  
GCGTGGCGGTGTACGCCATCTGCTTCATCCTAGCGGCCATCGCCATCCTGCTGAACCTGGGG  
GAGTGCACCAACGTGCTACCCATCCCCTTCCCCAGCTTCCTGTGGGGCTGGCCTTGCTGTC  
TGTCCTCCTCTATGCCACCGCCCTTGTTCTCTGGCCCCCTCTACCAGTTCGATGAGAAGTATG  
GCGGCCAGCCTCGGCGCTCGAGAGATGTAAGCTGCAGCCGCAGCCATGCCTACTACGTGTGT  
GCCTGGGACCGCCGACTGGCTGTGGCCATCCTGACGGCCATCAACCTACTGGCGTATGTGGC  
TGACCTGGTGCACCTCTGCCCACCTGGTTTTTTGTC AAGGTCTTAAAGACTCTCCCAAGAGGCTCC  
CGTTCCCTCTCCAACCTCTTTGTTCTTCTTGCCCGAGTTTTCTTTATGGAGTACTTCTTTCC  
TCCGCCCTTTCCTCTGTTTTTCTTCTTCTGCTCTCCCTCCCTCCACCTTTTTTCTTTCTTCC  
CAATTCCTTGCACTCTAACCAGTTCTTGGATGCATCTTCTTCCTTCCCTTTCCTCTTGCTGT  
TTCTTCTCTGTGTTGTTTTGTTGGCCACATCCTGTTTTTACCCCTGAGCTGTTTCTCTTTTT  
CTTTTCTTTCTTTTTTTTTTTTTTTTTTTTTTAAAGACGGATTCTCACTCTGTGGCCCAGGCTGGAG  
TGCAGTGGTGCATCTCAGCTCACTGCAACCCCCGCCTCCTGGGTTC AAGCGATTCTCCTCC  
CCCAGCCTCCCAAGTAGCTGGGAGGACAGGTGTGAGCTGCCGCACCCAGCCTGTTTCTCTTT  
TTCCACTCTTCTTTTTTCTCATCTCTTTTCTGGGTGTCCTGTGCGCTTTCTTATCTGCCTGT  
TTTGCAAGCACCTTCTCCTGTGTCTTGGGAGCCCTGAGACTTCTTTCTCTCCTTGCCCTCA  
CCCACCTCCAAAGGTGCTGAGCTCACATCCACACCCCTTGAGCCGTCCATGCCACAGCCCC  
CCAAGGGGGCCCCATTGCCAAAGCATGCCTGCCACCCCTCGCTGTGCCTTAGTCAGTGTGTAC  
GTGTGTGTGTGTGTGTGTTTGGGGGGTGGGGGGTGGGTAGCTGGGGATTGGGGCCCTCTTTCT  
CCCAGTGGAGGAAGGTGTGCAGTGTACTTCCCCTTTAAATTAAAAACATATATATATATAT  
ATTTGGAGGTCAGTAATTTCCAATGGGCGGGAGGCATTAAGCACCGACCTGGGTCCCTAGG  
CCCCGCCTGGCACTCAGCCTTGCCAGAGATTGGCTCCAGAATTTTTGCCAGGCTTACAGAACAC  
CCTACTGCCTAGAGGCCATCTTAAAGGAAGCAGGGGGCTGGATGCCTTTCATCCCAACTATTCT  
CTGTGGGTATGAAAAAG

**FIGURE 20**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58727  
<subunit 1 of 1, 322 aa, 1 stop  
<MW: 35274, pI: 8.57, NX(S/T): 1  
MPVTVTRTTITTTTTSSSGLGSPMIVGSPRALTQPLGLLRLQLVSTCVAFSLVASVGAWTG  
SMGNWSMFTWCFCFSVTLIILIVELCGLQARFPLSWRNFPI TFACYAALFCLSASIIYPTTY  
VQFLSHGRSRDHAI AATFFSCIACVAYATEVAWTRARPGEITGYMATVPGLLKVLETFVACI  
IFAFISDPNLYQHQP ALEWCVAVYAICFILAAIAILLNLGECTNVLPPIPFPSFLSGLALLSV  
LLYATALVLWPLYQFDEKYGGQPRRSRDVSCSRSHAYYVCAWDRRLAVAILTAINLLAYVAD  
LVHSAHLVFVKV

**Important features:****Transmembrane domains:**

amino acids 41-60 (type II), 66-85, 101-120, 137-153, 171-192,  
205-226, 235-255 and 294-312

**N-glycosylation site.**

amino acids 66-69

**Glycosaminoglycan attachment site.**

amino acids 18-21



[illegible]

**FIGURE 22**

MFLATLSFLLPFAHPFGTVSCEYMLGSPLSSLAQVNLSPFSHPKVHMDPNYCHPSTSLHLCS  
LAWSFTRLLHPPLSPGISQVVVDHVTKPTAMAQGRVAHLIEWKGWSKPSDSPAALSAFSSY  
SDLSEGEQEARFAAGVAEQFAIAEAKLRAWSSVDGEDSTDDSYDEDFAGGMDTDMAGQLPLG  
PHLQDLFTGHRFSRPVRQGSVEPESDCSQTVPDTLCSSSLCSLEDGLLGSPARLASQLLGDE  
LLLAKLPPSRESAFRSLGPLEAQDSLNSPLTESCLSPAEEEEPAPCKDCQPLCPPLTGSWER  
QRQASDLASSGVVSLDEDEAEPEEQ

**Signal peptide:**

amino acids 1-15

**Casein kinase II phosphorylation site.**

amino acids 123-127, 128-132, 155-159, 162-166, 166-170, 228-232,  
285-289, 324-328

**Tyrosine kinase phosphorylation site.**

amino acids 44-52

**N-myristoylation site.**

amino acids 17-23, 26-32, 173-179

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 11-22

**FIGURE 23**

GGTTCCTGGGCGCTCTGTTACACAAGCAAGATACAGCCAGCCCCACCTAATTTTGTTCCT  
GGCACCCCTCCTGCTCAGTGCACACTTAACCCATCTGTTTTCTCTAATGCACGA  
CAGATTCCTTTTAGACAGGACAACCTGTGATATTTAGTTCCCTGATTGTAAATACCTCCTAAG  
CCTGAAGCTTCTGTTACTAGCCATTGTGAGCTTCAGTTTCTTCATCTGCAAAATGGGCATAA  
TACAATCTATTCTTGCCACATCAAGGGATTGTTATTCCTTTAAAAAAAACCAATACCAAAG  
AAGCCTACA**ATG**TTGGCCTTAGCCAAAATCTGTTGATTTCAACGTTGTTTTATTCACTTCT  
ATCGGGGAGCCATGGAAAAGAAAATCAAGACATAAACACAACACAGAACATTGCAGAAGTTT  
TAAAACAATGGAAAATAAACCTATTTCTTTGGAAAGTGAAGCAAACCTAAACTCAGATAAA  
GAAAATATAACCACCTCAAATCTCAAGGCGAGTCATTCCCCCTCCTTTGAATCTACCCAACAA  
CAGCCACGGAATAACAGATTTCTCCAGTAACTCATCAGCAGAGCATTCTTTGGGCAGTCTAA  
AACCACATCTACCATTTCCACAAGCCCTCCCTTGATCCATAGCTTTGTTTCTAAAGTGCCT  
TGGAATGCACCTATAGCAGATGAAGATCTTTTGCCCATCTCAGCACATCCCAATGCTACACC  
TGCTCTGTCTTCAGAAAACCTTCACTTGGTCTTTGGTCAATGACACCGTGAAAACCTCTGATA  
ACAGTTCCATTACAGTTAGCATCCTCTCTTCAGAACCAACTTCTCCATCTGTGACCCCTTG  
ATAGTGGAACCAAGTGGATGGCTTACCACAAACAGTGATAGCTTCACTGGGTTTACCCCTTA  
TCAAGAAAAACAACCTCTACAGCCTACCTTAAAATTCACCAATAATTCAAACTCTTTCCAA  
ATACGTCAGATCCCCAAAAGAAAATAGAAATACAGGAATAGTATTGGGGCCATTTTAGGT  
GCTATTCTGGGTGTCTCATTGCTTACTCTTGTGGGCTACTTGTGTGTGGAAAAGGAAAAC  
GGATTCATTTTCCCATCGGCGACTTTATGACGACAGAAATGAACCAGTTCTGCGATTAGACA  
ATGCACCGGAACCTTATGATGTGAGTTTTGGGAATTCTAGCTACTACAATCCAACCTTGAAT  
GATTCAGCCATGCCAGAAAGTGAAGAAAATGCACGTGATGGCATTCCCTATGGATGACATACC  
TCCACTTCGTACTTCTGTAT**TAG**AACTAACAGCAAAAAGGCGTTAAACAGCAAGTGTCATCTA  
CATCCTAGCCTTTTGACAAATTCATCTTTCAAAGGTTACACAAAATTACTGTCACGTGGAT  
TTTGTCAAGGAGAATCATAAAAGCAGGAGACCAGTAGCAGAAATGTAGACAGGATGTATCAT  
CCAAAGGTTTTCTTTCTTACAATTTTTGGCCATCCTGAGGCATTTACTAAGTAGCCTTAATT  
TGTATTTTAGTAGTATTTTCTTAGTAGAAAATATTTGTGGAATCAGATAAAACTAAAAGATT  
TCACCATTACAGCCCTGCCTCATAACTAAATAATAAAAATTATTCCACCAAAAAATTCTAAA  
ACAATGAAGATGACTCTTTACTGCTCTGCCTGAAGCCCTAGTACCATAATTCAAGATTGCAT  
TTTCTTAAATGAAAATTGAAAGGGTGCTTTTTAAAGAAAATTTGACTTAAAGCTAAAAAGAG  
GACATAGCCCAGAGTTTCTGTTATTGGGAAATTGAGGCAATAGAAATGACAGACCTGTATTC  
TAGTACGTTATAATTTTCTAGATCAGCACACACATGATCAGCCCCTGAGTTATGAAGCTGA  
CAATGACTGCATTCAACGGGGCCATGGCAGGAAAGCTGACCCTACCCAGGAAAGTAATAGCT  
TCTTTAAAGTCTTCAAAGGTTTTGGGAATTTAACTTGTCTTAATATATCTTAGGCTTCAA  
TTATTTGGGTGCCTTAAAACTCAATGAGAATCATGGT

**FIGURE 24**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58732

><subunit 1 of 1, 334 aa, 1 stop

><MW: 36294, pI: 4.98, NX(S/T): 13

MLALAKILLISTLFYSLLSGSHGKENQDINTTQNIAEVFKTMENKPISLESEANLNSDKENI  
TTSNLKASHSPPLNLPNNSHGITDFSSNSSAEHSLGSLKPTSTISTSPPLIHSEFVSKVPWNA  
PIADEDLLPISAHPNATPALSSNFSTWSLVNDTVKTPDNSSITVSILSSEPTSPSVTPLIVE  
PSGWLTTNSDSFTGFTPYQEKTTLQPTLKFTNNSKLFNPNTSDPQKENRNTGIVFGAILGAIL  
GVSLTLVGYLLCGKRKTDSEFSHRRLYDDRNEPVLRLDNAPEPYDVSFGNSSYYNPTLND  
SA  
MPPESEENARDGIPMDDIPPLRTSV

**Signal peptide:**

amino acids 1-23

**Transmembrane domain:**

amino acids 235-262

**N-glycosylation site.**

amino acids 30-34, 61-65, 79-83, 90-94, 148-152, 155-159,  
163-167, 218-222, 225-229, 298-302, 307-311

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 25**

AACAGGATCTCCTCTTGCAGTCTGCAGCCCAGGACGCTGATTCCAGCAGCGCCTTACCGCGC  
AGCCCGAAGATTCACT**ATG**GTGAAAATCGCCTTCAATACCCCTACCGCCGTGCAAAAGGAGG  
AGGCGCGGCAAGACGTGGAGGCCCTCCTGAGCCGCACGGTCAGAACTCAGATACTGACCGGC  
AAGGAGCTCCGAGTTGCCACCCAGGAAAAAGAGGGGCTCCTCTGGGAGATGTATGCTTACTCT  
CTTAGGCCTTTCATTCATCTTGGCAGGACTTATTGTTGGTGGAGCCTGCATTTACAAGTACT  
TCATGCCCAAGAGCACCATTACCGTGGAGAGATGTGCTTTTTTTGATTCTGAGGATCCTGCA  
AATTCCCTTCGTGGAGGAGAGCCTAACTTCCTGCCTGTGACTGAGGAGGCTGACATTCGTGA  
GGATGACAACATTGCAATCATTGATGTGCCTGTCCCCAGTTTCTCTGATAGTGACCCTGCAG  
CAATTATTCATGACTTTGAAAAGGGAATGACTGCTTACCTGGACTTGTTGCTGGGGAACTGC  
TATCTGATGCCCCTCAATACTTCTATTGTTATGCCTCCAAAAAATCTGGTAGAGCTCTTTGG  
CAAAGTGGCGAGTGGCAGATATCTGCCTCAAAGTTATGTGGTTCGAGAAGACCTAGTTGCTG  
TGGAGGAAATTCGTGATGTTAGTAACCTTGGCATCTTTATTTACCAACTTTGCAATAACAGA  
AAGTCCTTCCGCCTTCGTGCGAGAGACCTCTTGCTGGGTTTCAACAAACGTGCCATTGATAA  
ATGCTGGAAGATTAGACACTTCCCCAACGAATTTATTGTTGAGACCAAGATCTGTCAAGAG**T**  
**AA**GAGGCAACAGATAGAGTGTCTTGGTAATAAGAAGTCAGAGATTTACAATATGACTTTAA  
CATTAAAGGTTTATGGGATACTCAAGATATTTACTCATGCATTTACTCTATTGCTTATGCTTT  
AAAAAAAGGAAAAAAAAAAAAAAAAACTACTAACCCTGCAAGCTCTTGTCAAATTTTAGTTTAAT  
TGGCATTGCTTGTTTTTTGAAACTGAAATTACATGAGTTTCATTTTTCTTTGCATTTATAG  
GGTTTAGATTTCTGAAAGCAGCATGAATATATCACCTAACATCCTGACAATAAATTCATCC  
GTTGTTTTTTTTGTTTGTGTTTTTTCTTTTCCTTTAAGTAAGCTCTTTATTCATCTTATG  
GTGGAGCAATTTTAAATTTGAAATATTTTAAATTGTTTTTGAACTTTTTGTGTAAAATATA  
TCAGATCTCAACATTGTTGGTTTCTTTGTTTTTCATTTTGTACAACCTTCTTGAATTTAGA  
AATTACATCTTTGCAGTTCTGTTAGGTGCTCTGTAATTAACCTGACTTATATGTGAACAATT  
TTCATGAGACAGTCATTTTAACTAATGCAGTGATTCTTTCTCACTACTATCTGTATTGTGG  
AATGCACAAAATTGTGTAGGTGCTGAATGCTGTAAGGAGTTTAGGTTGTATGAATTCTACAA  
CCCTATAATAAATTTTACTCTATACAAAAAAAAAAAAAAAAAAAAA

**FIGURE 26**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58828

<subunit 1 of 1, 263 aa, 1 stop

<MW: 29741, pI: 5.74, NX(S/T): 1

MVKIAFNTPTAVQKEEARQDVEALLSRTVVRTQILTGKELRVATQEKEGSSGRCMLTLLGLSF  
ILAGLIVGGACIYKYFMPKSTIYRGEMCFFDSEDPANSLRGGEPNFLPVTEEADIREDDNIA  
IIDVPVPSFSDSDPAAIHDFEKGMTAYLDLLLGNCYLMPLNTSIVMPPKNLVELFGKLASG  
RYLPQTYVVREDLVAVEEIRDVSNLGIFIYQLCNNRKSFRLRRRDLLLGFNKRAIDKCWKIR  
HFPNEFIVETKICQE

**Type II transmembrane domain:**

amino acids 53-75

**N-glycosylation site.**

amino acids 166-170

**Casein kinase II phosphorylation site.**

amino acids 35-39, 132-136, 134-138

**N-myristoylation site.**

amino acids 66-72, 103-109

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 63-74

**FIGURE 27**

GGAGGAGGGAGGGCGGGCAGGCGCCAGCCCAGAGCAGCCCCGGGCACCAGCACGGACTCTCT  
 CTTCCAGCCCAGGTGCCCCCCTCTCGCTCCATTTCGGCGGGAGCACCCAGTCCTGTACGCC  
 AAGGAACTGGTCCTGGGGGCACCATGGTTTCGGCGGCAGCCCCAGCCTCCTCATCCTTCTG  
 TTGCTGCTGCTGGGGTCTGTGCCTGCTACCGACGCCCGCTCTGTGCCCTGAAGGCCACGTT  
 CCTGGAGGATGTGGCGGGTAGTGGGGAGGCCGAGGGCTCGTCGGCCTCCTCCCCGAGCCTCC  
 CGCCACCCTGGACCCCGGCCCTCAGCCCCACATCGATGGGGCCCCAGCCCACAACCCTGGGG  
 GGCCCATCACCCCCACCAACTTCCTGGATGGGATAGTGGACTTCTTCCGCCAGTACGTGAT  
 GCTGATTGCTGTGGTGGGGCTCCCTGGCCTTTCTGCTGATGTTTCATCGTCTGTGCCGCGGTCA  
 TCACCCGGCAGAAGCAGAAGGCCTCGGCCTATTACCCATCGTCCTTCCCCAAGAAGAAGTAC  
 GTGGACCAGAGTGACCGGGCGGGGGCCCCCGGGCCTTCAGTGAGGTCCCCGACAGAGCCCC  
 CGACAGCAGGCCCCGAGGAAGCCCTGGATTTCCTCCCGGCAGCTCCAGGCCGACATCTTGGCCG  
 CCACCCAGAACCTCAAGTCCCCCACCAGGGCTGCACTGGGCGGTGGGGACGGAGCCAGGATG  
 GTGGAGGGCAGGGGCGCAGAGGAAGAGGAGAAGGGCAGCCAGGAGGGGGACCAGGAAGTCCA  
 GGGACATGGGGTCCCAGTGGAGACACCAGAGGCGCAGGAGGAGCCGTGCTCAGGGGTCTTG  
 AGGGGGCTGTGGTGGCCGGTGAGGGCCAAGGGGAGCTGGAAGGGTCTCTCTTGTAGCCCAG  
 GAAGCCCAGGGACCAGTGGGTCCCCCGAAAGCCCCTGTGCTTGCAGCAGTGTCCACCCCAG  
 TGTTAACAGTCCTCCCGGGCTGCCAGCCCTGACTGTGGGGCCCCCAAGTGGTCACCTCCCC  
 GTGTATGAAAAGGCCTTCAGCCCTGACTGCTTCCTGACACTCCCTCCTTGGCCTCCCTGTGG  
 TGCCAATCCCAGCATGTGCTGATTCTACAGCAGGCAGAAATGCTGGTCCCCGGTGCCCCGGA  
 GGAATCTTACCAAGTGCCATCATCCTTCACCTCAGCAGCCCCAAAGGGCTACATCCTACAGC  
 ACAGCTCCCCTGACAAAGTGAGGGAGGGCACGTGTCCCTGTGACAGCCAGGATAAAACATCC  
 CCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGTGCCCGGCCCAAACCTACTTTTTTAAACA  
 GCTACAGGGTAAAATCCTGCAGCACCCACTCTGGAAAATACTGCTCTTAATTTTCCTGAAGG  
 TGGCCCCCTGTTTCTAGTTGGTCCAGGATTAGGGATGTGGGGTATAGGGCATTTAAATCCTC  
 TCAAGCGCTCTCCAAGCACCCCCGGCCTGGGGGTGAGTTTCTCATCCCGCTACTGCTGCTGG  
 GATCAGGTTGAATGAATGGAACCTTCCTGTCTGGCCTCCAAAGCAGCCTAGAAGCTGAGGG  
 GCTGTGTTTGAGGGGACCTCCACCCTGGGGAAGTCCGAGGGGCTGGGGAAGGGTTTCTGACG  
 CCCAGCCTGGAGCAGGGGGGCCCTGGCCACCCCTGTTGCTCACACATTGTCTGGCAGCCTG  
 TGTCCACAATATTCGTCACTCCTCGACAGGGAGCCTGGGCTCCGTCCTGCTTTAGGGAGGCT  
 CTGGCAGGAGGTCTCTCCCCCATCCCTCCATCTGGGGCTCCCCAACCTCTGCACAGCTCT  
 CCAGGTGCTGAGATATAATGCACCAGCACATAAACCTTTATTCCGGCCTGAAAAAAAAAAAA  
 AAAGA

**FIGURE 28**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58852

><subunit 1 of 1, 283 aa, 1 stop

><MW: 29191, pI: 4.52, NX(S/T): 0

MVSAAAPSLILLILLLLLLGSVPATDARSVPLKATFLEDVAGSGEAEAGSSASSPSLPPPWTPAL  
SPTSMGPQPTTLGGPSPPTNFLDGIVDFFRQYVMLIAVVGSLAFLLMFIVCAAVITRQKQKA  
SAYYPSSFPPKKKYVDQSDRAGGPRAFSEVPDRAPDSRPEEALDSSRQLQADILAATQNLKSP  
TRAAALGGGDGARMVEGRGAEEEEKGSQEGDQEVQGHGVPVETPEAQEEPCSGVLEGAVVAGE  
GQGELEGSLLLAQEAQGPVGPPESPCACSSVHPSV

**Signal peptide:**

amino acids 1-25

**Transmembrane domain:**

amino acids 94-118

**N-myristoylation site.**

amino acids 18-24, 40-46, 46-52, 145-151, 192-198, 193-199,  
211-217, 238-244, 242-248



## FIGURE 29

[illegible]

**FIGURE 30**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59212
><subunit 1 of 1, 440 aa, 1 stop
><MW: 42208, pI: 6.36, NX(S/T): 1
MKFQGPLACLILLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGA
AGSKVSEALGQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHG
ADAVRGSWQGVPGHSGAWETSGGHGIFGSQGGGLGGQGNPGGLGTPWVHGYPGNSAGSFGM
NPQGAPWGQGGNGGPPNFGTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGS
GSQSGSSGSGSGNGDNNNGSSSSGSSSSGSSSSGSSSSGSSSSGSSSGNSGGSRGDSGSESSW
GSSTGSSSGNHGSGGGNGHKPGCEKPGNEARGSGESGIQGFRGQGVSSNMREISKEGNRL
GGSGDNYRGQGSWSGSGGDAVGGVNTVNSETPGMFNFDTFWKNFKSKLGFINWDAINKDQ
RSSRIP
```

**Signal peptide:**

amino acids 1-21

**N-glycosylation site.**

amino acids 265-269

**Glycosaminoglycan attachment site.**

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

**Casein kinase II phosphorylation site.**

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

**N-myristoylation site.**

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70,  
 74-80, 90-96, 96-102, 130-136, 140-146, 149-155, 152-158,  
 155-161, 159-165, 163-169, 178-184, 190-196, 194-200, 199-205,  
 218-224, 236-242, 238-244, 239-245, 240-246, 245-251, 246-252,  
 249-252, 253-259, 256-262, 266-272, 270-276, 271-277, 275-281,  
 279-285, 283-289, 284-290, 287-293, 288-294, 291-297, 292-298,  
 295-301, 298-304, 305-311, 311-317, 315-321, 319-325, 322-328,  
 323-329, 325-331, 343-349, 354-360, 356-362, 374-380, 381-387,  
 383-389, 387-393, 389-395, 395-401

**Cell attachment sequence.**

amino acids 301-304

**FIGURE 31**

GACCGGTCCCTCCGGTCTGGATGTGCGGACTCTGCTGCAGCGAGGGCTGCAGGCCCGCCGGGCGGTGCTCACCG  
 TGCCCTGGCTGGTGGAGTTTCTCTCCTTTGCTGACCATGTTGTTCCCTTGCTGGAATATTACCGGGACATCTTCA  
 CTCTCCTGCTGCGCCTGCACCGGAGCTTGGTGTGTGTCGAGGAGAGTGAGGGGAAGATGTTTCTGAACAAGC  
 TGCTGCTACTTGTGTCTGGGCTGGCTTTCCAGATTCCACAGTCCCTGAGGACTTGTCTTTCTGGAAGAGG  
 GTCCCTCATATGCCCTTGGAGGTGGACACAGTAGCCCCAGAGCATGGCTTGGACAATGCGCCTGTGGTGGACCAGC  
 AGCTGCTCTACACCTGCTGCCCCCTACATCGGAGAGCTCCGGAAACTGCTCGCTTCGTGGGTGTGAGGCAGTAGTG  
 GACGGAGTGGGGGCTTCATGAGGAAAATCACCCCCACCACTACCACCAGCCTGGGAGCCCAGCCTTCCCAGACCA  
 GCCAGGGGCTGCAGGCACAGCTCGCCAGGCCCTTTTCCACAACAGCCGCCCTCCTTGCGCCGAGCCGTAGAGT  
 TCGTGGCAGAAAGAATTGGATCAAACATGTGTCAAACATATCAAGGCTACACTGGTGGCAGATCTGGTGCGCCAGG  
 CAGAGTCACTTCTCCAAGAGCAGCTGGTGACACAGGGAGAGGAAGGGGGAGACCCAGCCCAGCTGTTGGAGATCT  
 TGTGTTCCCAGCTGTGCCCTCACGGGGCCCAGGCATTGGCCCTGGGGCGGGAGTTCTGTCAAAGGAAGAGCCCTG  
 GGGCTGTGCGGGCGCTGCTTCCAGAGGAGACCCCGGCAGCCGTTCTGAGCAGTGCAGAGAACATTGCTGTGGGGC  
 TTGCAACAGAGAAAAGCCTGTGCTTGGCTGTGAGCCAACATCACAGCACTGATCAGGAGGGAGGTGAAAGCAGCAG  
 TGAGTCGCACACTTCGAGCCCAGGGTCTGAACCTGCTGCCCCGGGGGAGCGGAGGGGCTGCTCCCGCGCCTGAC  
 GTGCTCTCCTTGGCCGTGGGGCCACGGGACCCTGACGAGGGAGTCTCCCAGAGCATCTGGAACAGCTCCTAGGC  
 CAGCTGGGCCAGACGCTGCGGTGCCGCCAGTTCTGTGCCCCACCTGCTGAGCAGCATCTGGCAAAGTGTCTGTG  
 GAGTTAGCTTCCCTCCTCGTTGCAGATCAAATTCCTATCCTAGGGCCCCCGGCACAGTACAGGCTGGAGAGAGGG  
 CAGGCTCGAAGGCTTCTGCACATGCTGCTTTCTTGTGGAAGGAAGACTTTTCAAGGGCCGTTCCGCTGCAGCTG  
 CTGCTGAGCCCAAGAAATGTGGGGCTTCTGGCAGACACAAGGCCAAGGGAGTGGGACTTGTGCTATTCTTGTCTA  
 CGGGAGCTGGTGGAGAAGGGTCTGATGGGACGGATGGAGATAGAGGCCTGCCTGGGCAGCCTCCACCAGGCCCCAG  
 TGGCCAGGGGACTTTGCTGAAGAATTAGCAACACTGTCTAATCTGTTTCTAGCCAGCCCCACCTGCCAGAACCC  
 CAGCTAAGAGCCTGTGAGTTGGTGCAGCCAAACCGGGGCACTGTGCTGGCCAGAGCTAGGGCTGAGAAGTGGCC  
 CTGCCTTGGGCATTGCACCAGAACCCTGGACCCCCGCTCACGAGGAGGCCCAAGTGCCCAATGCAGACCCTCAC  
 TGGTTGGGGTGTAGCTGGGTCTACAGTCAGACTTCTGCTCTAAGGGTGTCACTGCCTGGCATCCCACCACGCGA  
 ATCCTAGAGGAAGGAGAGTTGGCCTGATTTGGGATTATGGCAGAAAAGTCCAGAGATGCCAGTCCCTGGAGTAGAA  
 GAGGTGGTGTGTTTATCTCTTGGATACTAAATGAAATGAGGTGTGTGGGCTTGTCAACACAGAATTCAAGCCT  
 CATTTGCTATCCCAGCATCTCTTAAACCTTTGTAGTCTTGGAAATTCATGACAGAGGCAAATGACTCCTGCTTAAC  
 TTATGAAGAAAGTTAAACATGAATCTTGGGAGTCTACATTTTCTTATCACCAGGAGCTGGACTGCCATCTCCTT  
 ATAAATGCCTAACACAGGCCCGGTCTGGTGGCTCATGCCTGTAATCCCAGCACTTTGAGAGGCCTGAGGTGCGCG  
 GACTGCCTGAGGTGAGGAATTCAAGACCAGCCTGGCCAACATGGCAAACCCCATCTCTACTAAAAATAAAAAA  
 TTATTAGCTGGGCATGGTGGTGTGTGCTGTAATCCCAGCTACTCAGGAGGATGAGGCAGGAGACCTGCTTGAAC  
 CTGGAGGTGGAGGTGTCAGTGAGCCGAGGTGCGACCACTGCACTCCAGTCTGGGTAAACAGAGCGAGACTTTCTAG  
 AAAAAAGCCTAACAAACAGATAAGGTAGGACTCAACCAACTGAAACCTGACTTTCCCCCTGTACCTTCAGCCCCCTG  
 TGCAGGTAGTAACCTCTTGAGACCTCTCCCTGACCAGGGACCAAGCACAGGGCATTTAGAGCTTTTTAGAAATAAA  
 CTGGTTTTCTTTAAAAAAGGGCTTTTATTAATAATTCTCCACACAGATGGCTCCTGCAATCTGCCACAGCTC  
 TGGGGCGTGTCTGTAGGGAAAGGCCCTGTTTTCCCTGAGGCGGGGCTGGGCTTGTCCATGGGTCCGCGGAGCTG  
 GCCGTGCTTGGCGCCCTGGCGTGTGTCTAGCTGCTTGGCGGGCACAGAGCTGCGGGGTCTGGGGCACCGGG  
 AGCTAAGAGCAGGCTCTGGTGCAGGGGTGGAGGCCTGTCTCTTAACCGACACCTGAGGTGCTCCTGAGATGCTG  
 GGTCCACCCTGAGTGGCACGGGAGCAGCTGTGGCCGGTGTCTCTCYTAGGCCAGTCTTGGGGAAACTAAGCTC  
 GGGCCCTTCTTTGCAAAGACCGAGGATGGGGTGGGTGTGGGGGACTCATGGGGAATGGCCTGAGGAGCTACGTGT  
 GAAGAGGGCGCCGTTTGTGGCTGCAGCGGCCTGGAGCGCTCTCTCCTGAGCCTCAGTTTCCCTTTCCGTCTA  
 ATGAAGAACATGCCGTCTCGGTGTCTCAGGGCTATTAGGACTTGGCCCTCAGGAAGTGGCCTTGGACGAGCGTCAT  
 GTTATTTTCACAACTGCTTGCAGCGTTGGCCTGGGCACGTGATGGAATGGCCCATGTCCCTCTGCTGCGTGGAC  
 GTCGCGGTGCGGAGTGCGCAGCCAGAGGCGGGCCAGACGTGCGCCTGGGGGTGAGGGGAGGCGCCCCGGGAGGG  
 CCTCACAGGAAGTTGGGCTCCCGCACCAACAGGCAGGGCGGGCTCCCGCCCGCCGCCGCCACCACCGTCCAGG  
 GGCCGGTAGACAAAGTGGAAGTGCAGCTTGGGCTCGCTGCGCAGCAGGTAGCCCTTGATGCAGTGCGGCAGCGCG  
 TCGTCCGCCAGCTGGAAGCAGCGCCCGTCCACCAGCACGAACAGCCGGTGCACCT

**FIGURE 32**

MCFLNKL L L L L L A V L G W L F Q I P T V P E D L F F L E E G P S Y A F E V D T V A P E H G L D N A P V V D Q Q L L Y T C  
C P Y I G E L R K L L A S W S G S S G R S G G F M R K I T P T T T T S L G A Q P S Q T S Q G L Q A Q L A Q A F F H N Q P P  
S L R R T V E F V A E R I G S N C V K H I K A T L V A D L V R Q A E S L L Q E Q L V T Q G E E G D P A Q L L E I L C S Q L  
C P H G A Q A L A L G R E F C Q R K S P G A V R A L L P E E T P A A V L S S A E N I A V G L A T E K A C A W L S A N I T A L  
I R R E V K A A V S R T L R A Q G P E P A A R G E R R G C S R A

**Signal peptide:**

amino acids 1-18

**N-glycosylation site.**

amino acids 244-248

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 89-93

**Casein kinase II phosphorylation site.**

amino acids 21-25, 167-171, 223-227

**N-myristoylation site.**

amino acids 100-106, 172-178, 207-213

**Microbodies C-terminal targeting signal.**

amino acids 278-282

**FIGURE 33**

TCCCTTGACAGGTCTGGTGGCTGGTTTCGGGGTCTACTGAAGGCTGTCTTGATCAGGAAACTG  
 AAGACTCTCTGCTTTTGCCACAGCAGTTCCCTGCAGCTTCCTTGAGGTGTGAACCCACATCCC  
 TGCCCCCAGGGCCACCTGCAGGACGCCGACACCTACCCCTCAGCAGACGCCGGAGAGAA**ATG**  
 AGTAGCAACAAAGAGCAGCGGTGAGCAGTGTTCGTGATCCTCTTTGCCCTCATCACCATCCT  
 CATCCTCTACAGCTCCAACAGTGCCAATGAGGTCTTCCATTACGGCTCCCTGCGGGGCCGTA  
 GCCGCCGACCTGTCAACCTCAAGAAGTGGAGCATCACTGACGGCTATGTCCCCATTCTCGGC  
 AACAAGACACTGCCCTCTCGGTGCCACCAGTGTGTGATTGTCAGCAGCTCCAGCCACCTGCT  
 GGGCACCAAGCTGGGCCCTGAGATCGAGCGGGCTGAGTGTACAATCCGCATGAATGATGCAC  
 CCACCACTGGCTACTCAGCTGATGTGGGCAACAAGACCACCTACCGCGTCGTGGCCCCATTCC  
 AGTGTGTTCCGCGTGCTGAGGAGGCCCCAGGAGTTTGTCAACCGGACCCCTGAAACCGTGTT  
 CATCTTCTGGGGGGCCCCGAGCAAGATGCAGAAGCCCCAGGGCAGCCTCGTGCGTGTGATCC  
 AGCGAGCGGGCCTGGTGTTCGCCAACATGGAAGCATATGCCGTCTCTCCCGGCCGCATGCGG  
 CAATTTGACGACCTCTTCCGGGGTGAGACGGGCAAGGACAGGGAGAACTCTCATTCGTGGTT  
 GAGCACAGGCTGGTTTACCATGGTGATCGCGGTGGAGTTGTGTGACCACGTGCATGTCTATG  
 GCATGGTCCCCCCCCAACTACTGCAGCCAGCGGGCCCCGCCTCCAGCGCATGCCCTACCACTAC  
 TACGAGCCCAAGGGGCCGACGAATGTGTACCTACATCCAGAATGAGCACAGTCGCAAGGG  
 CAACCACCACCGCTTCATCACCCGAGAAAAGGGTCTTCTCATCGTGGGCCCAGCTGTATGGCA  
 TCACCTTCTCCCACCCCTCCTGGACCT**TAG**GCCACCCAGCCTGTGGGACCTCAGGAGGGTCAG  
 AGGAGAAGCAGCCTCCGCCCAGCCGCTAGGCCAGGGACCATCTTCTGGCCAATCAAGGCTTG  
 CTGGAGTGTCTCCAGCCAATCAGGGCCTTGAGGAGGATGTATCCTCCAGCCAATCAGGGCC  
 TGGGGAATCTGTTGGCGAATCAGGGATTTGGGAGTCTATGTGGTTAATCAGGGGTGTCTTTC  
 TTGTGCAGTCAGGGTCTGCGCACAGTCAATCAGGGTAGAGGGGTATTTCTGAGTCAATCTG  
 AGGCTAAGGACATGTCCTTTCCCATGAGGCCTTGGTTTCAGAGCCCCAGGAATGGACCCCCCA  
 ATCACTCCCCACTCTGCTGGGATAATGGGGTCCTGTCCCAAGGAGCTGGGAACCTTGGTGTTG  
 CCCCCTCAATTTCCAGCACCAGAAAGAGAGATTGTGTGGGGGTAGAAGCTGTCTGGAGGCCC  
 GGCCAGAGAATTTGTGGGGTTGTGGAGGTTGTGGGGGCGGTGGGGAGGTCCCAGAGGTGGGA  
 GGCTGGCATCCAGGTCTTGGCTCTGCCCTGAGACCTTGGACAAACCCTTCCCCCTCTCTGGG  
 CACCCTTCTGCCCCACACAGTTTCCAGTGCGGAGTCTGAGACCTTTCCACCTCCCCTACAA  
 GTGCCCTCGGGTCTGTCTCCCCGTCTGGACCTTCCCAGCCACTATCCCTTGCTGGAAGGCT  
 CAGCTCTTTGGGGGGTCTGGGGTGACCTCCCCACCTCCTGGAAAACCTTTAGGGTATTTTTGC  
 GCAAACCTCCTTCAGGGTTGGGGGACTCTGAAGGAAACGGGACAAAACCTTAAGCTGTTTTCT  
 TAGCCCCCTCAGCCAGCTGCCATTAGCTTGGCTCTTAAAGGGCCAGGCCTCCTTTTCTGCCCT  
 CTAGCAGGGAGGTTTTCCAACCTGTTGGAGGCGCCTTTGGGGCTGCCCTTTGTCTGGAGTCA  
 CTGGGGGCTTCCGAGGGTCTCCCTCGACCTCTGTTCCTGGGATGGCTGTGCGGAGCTGT  
 ATCACCTGGGTTCTGTCCCCCTGGCTCTGTATCAGGCACCTTTATTAAAGCTGGGCCTCAGTGG  
 GGTGTGTTTGTCTCCTGCTCTTCTGGAGCCTGGAAGGAAAGGGCTTCAGGAGGAGGCTGTGA  
 GGCTGGAGGGACCAGATGGAGGAGGCCAGCAGCTAGCCATTGCACACTGGGGTGATGGGTGG  
 GGGCGGTGACTGCCCCAGACTTGGTTTTGTAATGATTTGTACAGGAATAAACACACCTACGC  
 TCCGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 34**

MSSNKEQRSAVFVILFALITILILYSSNSANEVFHYGSLRGRSRRPVNLKKWSITDGYVPIL  
GNKTLPSRCHQCVIVSSSSHLLGTKLGPEIERAECTIRMNDAPTTGYSADVGNKTTYRVVAH  
SSVFRVLRRPQEFVNRTPETVFIFWGPSPKMQKPQGSLVRVIQRAGLVFPNMEAYAVSPGRM  
RQFDDLFRGETGKDREKSHSWLSTGWFTMVIAVELCDHVHVYGMVPPNYCSQRPRLQRMPYH  
YYEPKGPDECVTYIQNEHSRKGNNHHRFITEKRVFSSWAQLYGITFSHPSWT

**Signal peptide:**

amino acids 1-29

**Transmembrane domain:**

amino acids 9-31 (type II)

**N-glycosylation site.**

amino acids 64-68, 115-119

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 50-54

**Casein kinase II phosphorylation site.**

amino acids 3-7, 29-33, 53-57, 197-201

**Tyrosine kinase phosphorylation site.**

amino acids 253-262

**N-myristoylation site.**

amino acids 37-43, 114-120, 290-294



**FIGURE 36**

MLRGTMTAWRGMRPEVTLACLLLATAGCFADLNEVPQVTVQPASTVQKPGGTVILGCVVEPP  
 RMNVTWRLNGKELNGSDDALGVLITHGTLVITALNNHTVGRYQCVARMPAGAVASVPATVTL  
 ANLQDFKLDVQHVIEVDEGNTAVIACHLPESHKPAQVRYSVKQEWLEASRGNYLIMPSTGNLQ  
 IVNASQEDEGMYKCAAYNPVTQEVKTSGSSDRLRVRRSTAEAAARIIPPEAQTTIIVTKGQSL  
 ILECVASGIPPPRVTWAKDGSSVTGYNKTRFLLSNLLIDTTSEEDSGTYRCMADNGVGQPGA  
 AVILYNVQVFEPPEVTMELSOLVIPWGQSAKLTCEVRGNPPPSVLWLRNAVPLISSQRLRLS  
 RRALRVLSMGPEDEGVYQCAENEVGSAHAVVQLRTSRPSITPRLWQDAELATGTPPVSPSK  
 LGNPEQMLRGQPALPRPPTSVGPASPKCPGEGKGQAPAEAPIILSSPRTSKTDSYELVWRPR  
 HEGSGRAPILYYVVKHRKQVTNSSDDWTISGIPANQHRLTLRLDPGSLYEVEMAAYNCAGE  
 GQTAMVTFRTGRRPKPEIMASKEQQIQRDDPGASQSSSQPDHGRLSPPEAPDRPTISTASE  
 TSVYVTWIPRGNGGFPIQSFRVEYKKLKKVGDWILATSAIPPSRLSVEITGLEKGTSYKFRV  
 RALNMLGESEPSAPSRPYVVSIGYSGRVYERPVAGPYITFTDAVNETTIMLKWMIIPASNNNT  
 PIHGFYIYYRPTDSDNDSYKKDMVEGDKYWHSISHLQPETSIDIKMQCFNEGGESEFSNVM  
 ICETKARKSSGQPGRLPPPTLAPPQPPLPETIERPVGTTAMVARSSDLPLYLIVGVVLGSIVL  
 IIVTFIPFCLWRAWKQKHTTDLGFPRSALPPSCPYTMVPLGGLPGHQASGQPYLSGISGRA  
 CANGIHMNRGCPSAAVGYPGMKPQQHCPGELQQQSDTSSLLRQTHLGNGYDPQSHQITRGPK  
 SSPDEGSFLYTLPPDDSTHQLLQPHHDCCQRQEQAAGVQSGVRRAPDSPVLEAVWDPPFHSG  
 PPCCCLGLVPVEEVDSPDSCQVSGGDWCPQHPVGAYVGQEPGMQLSPGPLVRVSFETPPLTI

**Signal peptide:**

amino acids 1-30

**Transmembrane domain:**

amino acids 16-30 (type II), 854-879



**FIGURE 37**

CGGGAGGCTGGGTCGTCATGATCCGGACCCCATTTGTCGGCCTCTGCCCATCGCCTGCTCCTC  
CCAGGCTCCCGCGGCCGACCCCGCGCAAC**ATG**CAGCCACGGGCCGCGAGGGTTCCCGCGC  
GCTCAGCCGGCGGTATCTGCGGCGTCTGCTGCTCCTGCTACTGCTGCTGCTGCTGCGGCAGC  
CCGTAACCCGCGCGGAGACCACGCCGGGCGCCCCCAGAGCCCTCTCCACGCTGGGCTCCCC  
AGCCTCTTCACCACGCCGGGTGTCCCCAGCGCCCTCACTACCCAGGCCTCACTACGCCAGG  
CACCCCAAAACCCTGGACCTTCGGGGTTCGCGCGCAGGCCCTGATGCGGAGTTTCCCACTCG  
TGGACGGCCACAATGACCTGCCCCAGGTCTGAGACAGCGTTACAAGAATGTGCTTCAGGAT  
GTTAACCTGCGAAATTTTCAGCCATGGTCAGACCAGCCTGGACAGGCTTAGAGACGGCCTCGT  
GGGTGCCAGTTCTGGTCAGCCTCCGTCTCATGCCAGTCCCAGGACCAGACTGCCGTGCGCC  
TCGCCCTGGAGCAGATTGACCTCATTCACCGCATGTGTGCCTCCTACTCTGAACTCGAGCTT  
GTGACCTCAGCTGAAGGTCTGAACAGCTCTCAAAGCTGGCCTGCCTCATTGGCGTGNAGGG  
TGGTCACTCACTGGACAGCAGCCTCTCTGTGCTGCGCAGTTTCTATGTGCTGGGGGTGCGCT  
ACCTGACACTTACCTTCACCTGCAGTACCCATGGGCAGAGAGTTCCACCAAGTTCAGACAC  
CACATGTACACCAACGTCAGCGGATTGACAAGCTTTGGTGAGAAAGTAGTAGAGGAGTTGAA  
CCGCCTGGGCATGATGATAGATTTGTCTATGCATCGGACACCTTGATAAGAAGGGTCCTGG  
AAGTGTCTCAGGCTCCTGTGATCTTCTCCCACTCAGCTGCCAGAGCTGTGTGTGACAATTTG  
TTGAATGTTCCCGATGATATCCTGCAGCTTCTGAAGAACGGTGGCATCGTGATGGTGACACT  
GTCCATGGGGGTGCTGCAGTGCAACCTGCTTGCTAACGTGTCCACTGTGGCAGATCACTTTG  
ACCACATCAGGGCAGTCATTGGATCTGAGTTCATCGGGATTGGTGGAAATTATGACGGGACT  
GGCCGGTTCCTCAGGGGCTGGAGGATGTGTCCACATACCCAGTCCTGATAGAGGAGTTGCT  
GAGTCGTASCTGGAGCGAGGAAGAGCTTCAAGGTGTCCTTCGTGGAAACCTGCTGCGGGTCT  
TCAGACAAGTGGAAGGTGAGAGAGGAGAGCAGGGCGCAGAGCCCCGTGGAGGCTGAGTTT  
CCATATGGGCAACTGAGCACATCCTGCCACTCCCACCTCGTGCCTCAGAATGGACACCAGGC  
TACTCATCTGGAGGTGACCAAGCAGCCAACCAATCGGGTCCCCTGGAGGTCCCTCAAATGCCT  
CCCCATACCTTGTTCCAGGCCTTGTTGGCTGCTGCCACCATCCCAACCTTCACCCAGTGGCTC  
TGCT**TGA**CACAGTCGGTCCCCGCAGAGGTCACTGTGGCAAAGCCTCACAAAGCCCCCTCTCCT  
AGTTCATTACAAAGCATATGCTGAGAATAAACATGTTACACATGGAAAA

**FIGURE 38**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59817

><subunit 1 of 1, 487 aa, 1 stop, 2 unknown

><MW: 53569.32, pI: 7.68, NX(S/T): 5

MQPTGREGSRALSRRYLRRLLLLLLLLLLLRQPVTRAETTPGAPRALSTLGSPSLFTTPGVPS  
 ALTTPGLTTPGTPKTLDLRGRAQALMRSFPLVDGHNDLPQVLRQRYKNVLQDVNLRNFSHGQ  
 TSLDRLRDGLVGAQFWSASVSCQSQDQTAVRLALEQIDLIHRMCASYSELELVTSAEGLNSS  
 QKLACLIGVXGGHSLDSSLSVLRSFYVLGVRYLTLTFTCSTPWAESSTKFRHHMYTNVSGLT  
 SFGEKVVEELNRLGMMIDLSYASDTLIRRVLEVSQAPVIFSHSAARAVCDNLLNVPDDILQL  
 LKNGGIVMVTLSMGVLQC�LLANVSTVADHFDHIRAVIGSEFIGIGGNYDGTGRFPQGLEDV  
 STYPVLIEELLRSXWSEEEELQGVLRGNLLRVFRQVEKVREESRAQSPVEAEFPYQQLSTSCH  
 SHLVPQNGHQATHLEVTKQPTNRVPWRSSNASPYLVPGLVAAATIPTFTQWLC

**Important features of the protein:**

**Signal peptide:**

amino acids 1-36

**Transmembrane domain:**

amino acids 313-331

**N-glycosylation sites.**

amino acids 119-122, 184-187, 243-246 and 333-336

**N-myristoylation sites.**

amino acids 41-46, 59-64, 73-78, 133-138, 182-187, 194-199, 324-329, 354-359, 357-362, 394-399, 427-432 and 472-477.

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 136-146

**FIGURE 39**

TGCTAGGCTCTGTCCCACAATGCACCCGAGAGCAGGAGCTGAAAGCCTCTAACACCCACAGA  
 TCCCTCTATGACTGCAATGTGAGGTGTCCGGCTTTGCTGGCCCAGCAAGCCTGATAAGC**ATG**  
 AAGCTCTTATCTTTGGTGGCTGTGGTCGGGTGTTTGCTGGTGCCCCCAGCTGAAGCCAACAA  
 GAGTTCTGAAGATATCCGGTGCAAATGCATCTGTCCACCTTATAGAAACATCAGTGGGCACA  
 TTTACAACCAGAATGTATCCCAGAAGGACTGCAACTGCCTGCACGTGGTGGAGCCCATGCCA  
 GTGCCTGGCCATGACGTGGAGGCCTACTGCCTGCTGTGCGAGTGCAGGTACGAGGAGCGCAG  
 CACCACCACCATCAAGGTCATCATTGTCATCTACCTGTCCGTGGTGGGTGCCCTGTTGCTCT  
 ACATGGCCTTCCTGATGCTGGTGGACCCTCTGATCCGAAAGCCGGATGCATACACTGAGCAA  
 CTGCACAATGAGGAGGAGAATGAGGATGCTCGCTCTATGGCAGCAGCTGCTGCATCCCTCGG  
 GGGACCCCGAGCAAACACAGTCCTGGAGCGTGTGGAAGGTGCCCAGCAGCGGTGGAAGCTGC  
 AAGGTGCAGGAGCAGCGGAAGACAGTCTTCGATCGGCACAAGATGCTCAGC**TAG**ATGGGCTGG  
 TGTGGTTGGGTCAAGGCCCCAACACCATGGCTGCCAGCTTCCAGGCTGGACAAAGCAGGGGG  
 CTACTTCTCCCTTCCCTCGGTTCCAGTCTTCCCTTTAAAGCCTGTGGCATTTTTCTCCTT  
 CTCCCTAACTTTAGAAATGTTGTACTTGGCTATTTTGATTAGGGAAGAGGGATGTGGTCTCT  
 GATCTCTGTTGTCTTCTTGGGTCTTTGGGGTTGAAGGGAGGGGGAAGGCAGGCCAGAAGGGA  
 ATGGAGACATTCGAGGCGGCCTCAGGAGTGGATGCGATCTGTCTCTCCTGGCTCCACTCTTG  
 CCGCCTTCCAGCTCTGAGTCTTGGGAATGTTGTTACCCTTGAAGATAAAGCTGGGTCTTCA  
 GGAACCTCAGTGTCTGGGAGGAAAGCATGGCCCAGCATTGAGCATGTGTTCCCTTCTGCAGTG  
 GTTCTTATCACCACCTCCCTCCCAGCCCCGGCGCCTCAGCCCCAGCCCCAGCTCCAGCCCTG  
 AGGACAGCTCTGATGGGAGAGCTGGGCCCCCTGAGCCCACTGGGTCTTCAGGGTGCACCTGGA  
 AGCTGGTGTTCGCTGTCCCCTGTGCACTTCTCGCACTGGGGCATGGAGTGCCCATGCATACT  
 CTGCTGCCGGTCCCCTCACCTGCACTTGAGGGGTCTGGGCAGTCCCTCCTCTCCCCAGTGTC  
 CACAGTCACTGAGCCAGACGGTCGGTTGGAACATGAGACTCGAGGCTGAGCGTGGATCTGAA  
 CACCACAGCCCCGTGACTTGGGTGCCTCTTGTCCCTGAACTTCGTTGTACCAGTGCATGGA  
 GAGAAAATTTTGTCTCTTGTCTTAGAGTTGTGTGTAAATCAAGGAAGCCATCATTAAATTG  
 TTTTATTTCTCTCA

**FIGURE 40**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60278

<subunit 1 of 1, 183 aa, 1 stop

<MW: 20574, pI: 6.60, NX(S/T): 3

MKLLSLVAVVGCLLVPPAEANKSSEDIRCKCICPPYRNISGHIYNQNVSQKDCNCLHVVEPM  
PVPBGHDVEAYCLLCECRYEERSTTTIKVIIIVYLSVVGALLLYMAFLMLVDPLIRKPDAYTE  
QLHNEEENEDARSMAAAAASLGGPRANTVLERVEGAQQRWKLQVQEQRKTVFDRHKMLS

**Important features:****Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 90-112

**N-glycosylation sites.**

amino acids 21-24, 38-41 and 47-50

**FIGURE 41**

AGCGGGTCTCGCTTGGGTTCCGCTAATTTCTGTCTGAGGCGTGAGACTGAGTTCATAGGGTCCTGGGTCCCCGA  
 ACCAGGAAGGGTTGAGGGAACACAATCTGCAAGCCCCCGCAGCCCAAGTGAGGGGCCCCGTGTTGGGGTCCTCCC  
 TCCCTTTGCATTCCCACCCCTCCGGGCTTTGCGTCTTCTGGGGACCCCTCGCCGGGAGATGGCCGCGTTGATG  
 CGGAGCAAGGATTCGTCTGCTGCCTGCTCCTACTGGCCGCGGTGCTGATGGTGGAGAGCTCACAGATCGGCAGT  
 TCGCGGGCCAACTCACTCCATCAAGTCTCTCTGGGCGGGGAGACGCCTGGTCAGGCCGCCAATCGATCTGCG  
 GGCATGTACCAAGGACTGGCATTCTGGCGGCAGTAAGAAGGGCAAAAACCTGGGGCAGGCCTACCTTTGTAGCAGT  
 GATAAGGAGTGTGAAGTTGGGAGGTATTGCCACAGTCCCCACCAAGGATCATCGGCCTGCATGGTGTGTCGGAGA  
 AAAAGAAGCGCTGCCACCGAGATGGCATGTGCTGCCCCAGTACCCGCTGCAATAATGGCATCTGTATCCCAGTT  
 ACTGAAAGCATCTTAACCCCTCACATCCCGGCTCTGGATGGTACTCGGCACAGAGATCGAAACCACGGTCATTAC  
 TCAAACCATGACTTGGGATGGCAGAATCTAGGAAGACCACACACTAAGATGTCACATATAAAAGGGCATGAAGGA  
 GACCCCTGCCTACGATCATCAGACTGCATTGAAGGGTTTTGCTGTGCTCGTCATTTCTGGACCAAAATCTGCAAA  
 CCAGTGCTCCATCAGGGGGAAGTCTGTACCAAAACAACGCAAGAAGGGTTCTCATGGGCTGGAAATTTTCCAGCGT  
 TCGGACTGTGCGAAGGGCTGTCTTGCAAAGTATGGAAGATGCCACCTACTCCTCCAAAGCCAGACTCCATGTG  
 TGTGAGAAAATTTGATCACCATTGAGGAACATCATCAATTGCAGACTGTGAAGTTGTGTATTTAATGCATTATAG  
 CATGGTGGAAAATAAGGTTTCAGATGCAGAAGAATGGCTAAAATAAGAAACGTGATAAGAATATAGATGATCACA  
 AAAGGGAGAAAAGAAAACATGAAGTGAATAGATTAGAATGGGTGACAAATGCAGTGCAGCCAGTGTTCATTATG  
 CACTTGTCTATGTAAATAATGTACACATTTGTGGAAAATGCTATTATTAAGAGAACAAGCACACAGTGGAAATT  
 ACTGATGAGTAGCATGTGACTTTCCAAGAGTTTAGGTTGTGCTGGAGGAGAGGTTTCCTTCAGATTGCTGATTGC  
 TTATACAAAATAACCTACATGCCAGATTTCTATTCAACGTTAGAGTTTAACAAAATACTCCTAGAATAACTTGTTA  
 TACAATAGGTTCTAAAAATAAAATTTGCTAAACAAGAAATGAAAACATGGAGCATTGTTAATTTACAACAGAAAA  
 TACCTTTTGATTGTAACTACTTCTGCTGTTCAATCAAGAGTCTTGGTAGATAAGAAAAAATCAGTCAATAT  
 TTCCAAATAATTGCAAAATAATGGCCAGTTGTTTAGGAAGGCCTTTAGGAAGACAAATAAATAACAAACAAACAG  
 CCACAAATACTTTTTTTTCAAAATTTTAGTTTTACCTGTAATTAATAAGAACTGATACAAGACAAAAACAGTTCC  
 TTCAGATTCTACGGAATGACAGTATATCTCTTTATCCTATGTGATTCTGCTCTGAATGCATTATATTTTCCA  
 AACTATACCCATAAATTTGACTAGTAAATACTTACACAGAGCAGAATTTTACAGATGGCAAAAAAATTTAA  
 GATGTCCAATATATGTGGGAAAAGAGCTAACAGAGAGATCATTATTTCTTAAAGATTGGCCATAACCTATATTTT  
 GATAGAATTAGATTGGTAAATACATGTATTCTATACATACTCTGTGGTAATAGAGACTTAAGCTGGATCTGTACTG  
 CACTGGAGTAAGCAAGAAAATTTGGGAAAACCTTTTTCGTTTGTTCAGGTTTTGGCAACACATAGATCATATGTCTG  
 AGGCACAAGTTGGCTGTTTCATCTTTGAAACCAGGGGATGCACAGTCTAAATGAATATCTGCATGGGATTTGCTAT  
 CATAATATTTACTATGCAGATGAATTCAGTGTGAGGTCCTGTGTCCGTACTATCCTCAAATTTATTTTATAG  
 TGCTGAGATCCTCAAATAATCTCAATTTAGGAGGTTTCACAAAATGTACTCCTGAAGTAGACAGAGTAGTGAGG  
 TTTTCATTGCCCTCTATAAGCTTCTGACTAGCCAATGGCATCATCCAATTTTCTTCCCAAACCTCTGCAGCATCTG  
 CTTTATTGCCAAAGGGCTAGTTTTCGGTTTTCTGCAGCCATTGCGGTTAAAAAATATAAGTAGGATAACTTGTA  
 AACCTGCATATTGCTAATCTATAGACACCACAGTTTCTAAATTTCTTGAACCACTTTACTACTTTTTTAACTT  
 AACTCAGTTCTAAATACTTTGTCTGGAGCACAAAACAATAAAAGGTTATCTTATAGTCGTGACTTTAACTTTTG  
 TAGACCACAATTCACTTTTAGTTTTCTTTTACTTAAATCCCATCTGCAGTCTCAAATTTAAGTTCTCCAGTAG  
 AGATTGAGTTTGAGCCTGTATATCTATTAATAATTTCACTTCCACATATATTTACTAAGATGATTAAGACTTA  
 CATTTTCTGCACAGGTCTGCAAAAACAAAAATTATAAACTAGTCCATCCAAGAACCAGTTTGTATAAACAGGT  
 TGCTATAAGCTTGTGAAATGAAATGGAACATTTCAATCAAACATTTCTATATAACAATTATTATATTACAAAT  
 TTGGTTTCTGCAATATTTTTCTTATGTCCACCCTTTTAAAAATTATTATTTGAAGTAATTTATTTACAGGAAATG  
 TTAATGAGATGTATTTCTTATAGAGATATTTCTTACAGAAAGCTTTGTAGCAGAATATATTTGCAGCTATTGAC  
 TTTGTAATTTAGGAAAAATGTATAATAAGATAAAATCTATTAAATTTTTCTCCTCTAAAAACTGAAAAAAAAA  
 AAAAAAAAAAAAAAAAAA

**FIGURE 42**

MAALMRSKDSSCCLLLLA AVL MV E S S Q I G S S R A K L N S I K S S L G G E T P G Q A A N R S A G M Y Q G L A  
F G G S K K G K N L G Q A Y P C S S D K E C E V G R Y C H S P H Q G S S A C M V C R R K K K R C H R D G M C C P S T R C N N  
G I C I P V T E S I L T P H I P A L D G T R H R D R N H G H Y S N H D L G W Q N L G R P H T K M S H I K G H E G D P C L R S  
S D C I E G F C C A R H F W T K I C K P V L H Q G E V C T K Q R K K G S H G L E I F Q R C D C A K G L S C K V W K D A T Y S  
S K A R L H V C Q K I

**Signal peptide:**

amino acids 1-25

bioRxiv preprint doi: <https://doi.org/10.1101/033300>; this version posted October 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 43**

TCTCAATCTGCTGACCTCGTGATCCGCCTGACCTTGTAATCCACCTACCTTGGCCTCCCAAA  
 GTGTTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCAACATCACGTTTTTAAAAATTGATT  
 TCTTCAAATTCATGGCAAATATTTCCCTTCCCTTTAACTTCTTATGTCAGAATGAGGAAGGA  
 TAGCTGCATTTATTTAGTCAGTTTTTCATTGCATAGTAATATTTTCATGTAGTATTTTCTAAG  
 TTATATTTTAGTAATTCATATGTTTTAGATTATAGGTTTTAACATACTTGTGAAAATACTTG  
ATGTGTTTTTAAAGCCTTGGGCAGAAATTCTGTATTGTTGAGGATTTGTTCTTTTATCCCCCT  
 TTTAAAGTCATCCGTCCTTGGCTCAGGATTTGGAGAGCTTGCACCACCAAAAATGGCAAACA  
 TCACCAGCTCCCAGATTTTGGACCAGTTGAAAGCTCCGAGTTTGGGCCAGTTTACCACCACC  
 CCAAGTACACAGCAGAATAGTACAAGTCACCTACAATACTACTTCTTGGGACCTCAAGCC  
 CCCAACATCCCAGTCCTCAGTCCTCAGTCATCTTGACTTCAAATCTCAACCTGAGCCATCCC  
 CAGTTCTTAGCCAGTTGAGCCAGCGACAACAGCACCAGAGCCAGGCAGTCACTGTTCCCTCCT  
 CCTGGTTTGGAGTCCTTTCCCTCCCAGGCCAAAACCTTCGAGAATCAACACCTGGAGACAGTCC  
 CTCCACTGTGAACAAGCTTTTGCAGCTTCCCAGCACGACCATTGAAAATATCTCTGTGTCTG  
 TCCACCAGCCACAGCCCAAACACATCAAACCTTGCTAAGCGGCGGATACCCCCAGCTTCTAAG  
 ATCCCAGCTTCTGCAGTGGAAATGCCTGGTTCAGCAGATGTCACAGGATTAAATGTGCAGTT  
 TGGGGCTCTGGAATTTGGGTCAGAACCTTCTCTCTCTGAATTTGGATCAGCTCCAAGCAGTG  
 AAAATAGTAATCAGATTCCCATCAGCTTGTATTTCGAAGTCTTTAAGTGAGCCTTTGAATACA  
 TCTTTATCAATGACCAGTGCAGTACAGAACTCCACATATACAACCTCCGTCATTACCTCCTG  
 CAGTCTGACAAGCTCATCACTGAATTCTGCTAGTCCAGTAGCAATGTCTTCCCTCTTATGACC  
 AGAGTTCTGTGCATAACAGGATCCCATACCAAAGCCCTGTGAGTTCATCAGAGTCAGCTCCA  
 GGAACCATCATGAATGGACATGGTGGTGGTTCGAAGTCAGCAGACACTAGACAGTAAGTATAG  
 CAGCAAGCTACTCTTGTGCATGGCTGGTGCCAACCAAACAGAGGAAGAGGATAGCTCACGTGA  
 TGTGGAACACACCAGTTGGTCAATGGCTCATTCGTTAAAAAGCAGCCCTTTTGCTTTTTTGT  
 TTTTGGACCAGGTGTTGGCTGTGGTGTATTAGAAATGTCTTAACCACAGCAAGAAGGAGGT  
 GGTGGTCTCATATTCTTCTGCCCTAATCAGACTGCACCACAAGTGCAGCATAACAGTATGCAT  
 TTTAAAGATGCTTGGGCCAGGCGGGGTGGCTGATGCCCATAAATCCCAGTGCTTTGGGGGGCC  
 AAGGCAGGCAGATTGCCCAAGCTCAGGAGTTTGAGACCACCCTGGGCAACATGGTGAAACTC  
 TGTCTCTACTAAAATACGAAAACTAGCCGGGTGTGGTGGCGGCGCGTGCCTGTAATCCCAG  
 CTACTTGGGAGGCTGAGGCACAAGAATCGCTTGAGCCAGCTTGGGCTACAAAGTGAGACTCC  
 GTCTGAAAAGA

**FIGURE 44**

MCFKALGRNSVLLRICSFIPLLKSSVLGSGFGELAPPKMANITSSQILDQLKAPSLGQFTTT  
PSTQQNSTSHPTTTTTSWDLKPPTSQSSVLSHLDFKSQPEPSPVLSQLSQRQQHQSQAVTVPP  
PGLESFPSQAKLRESTPGDSPSTVNKLLQLPSTTIENISVSVHQPPKHIKLAKRRIPPASK  
IPASAVEMPGSADVTGLNVQFGALEFGSEPSLSEFGSAPSSSENSNQIPISLYSKSLSEPLNT  
SLSMTSAVQNSTYTTTSVITSCSLTSSSLNSASPVAMSSSYDQSSVHNRIPTYQSPVSSSESAP  
GTIMNGHGGGRSQQTLDISKYSSKLLLSWLVP TKQRKRIAHVMWKTPVGQWLIR

**Signal peptide:**

amino acids 1-24

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100



**FIGURE 45**

GCCGAGTGGGACAAAGCCTGGGGCTGGGCGGGGGCC**ATG**GCGCTGCCATCCCGAATCCTGCT  
 TTGGAAACTTGTGCTTCTGCAGAGCTCTGCTGTTCTCCTGCACTCAGCGGTGGAGGAGACGG  
 ACGCGGGGCTGTACACCTGCAACCTGCACCATCACTACTGCCACCTCTACGAGAGCCTGGCC  
 GTCCGCCTGGAGGTACCGACGGCCCCCGGCCACCCCCGCTACTGGGACGGCGAGAAGGA  
 GGTGCTGGCGGTGGCGCGCGGCGCACCCGCGCTTCTGACCTGCGTGAACCGCGGGCACGTGT  
 GGACCGACCGGCACGTGGAGGAGGCTCAACAGGTGGTGCCTGGGACCGGCAGCCGCCCGGG  
 GTCCCGCACGACCGCGCGGACCGCCTGCTGGACCTCTACGCGTCGGGCGAGCGCCGCGCCTA  
 CGGGCCCCCTTTTTCTGCGCGACCGCGTGGCTGTGGGCGCGGATGCCTTTGAGCGCGGTGACT  
 TCTCACTGCGTATCGAGCCGCTGGAGGTGCGCGACGAGGGCACCTACTCCTGCCACCTGCAC  
 CACCATTACTGTGGCCTGCACGAACGCCGCGTCTTCCACCTGACGGTCGCCGAACCCACGC  
 GGAGCCGCCCCCCCGGGGCTCTCCGGGCAACGGCTCCAGCCACAGCGGCGCCCCAGGCCAG  
 ACCCCACACTGGCGCGCGGCCACAACGTCATCAATGTCATCGTCCCCGAGAGCCGAGCCAC  
 TTCTTCCAGCAGCTGGGCTACGTGCTGGCCACGCTGCTGCTCTTCATCCTGCTACTGGTCAC  
 TGTCTCCTGGCCGCCCGCAGGCGCCGCGGAGGCTACGAATACTCGGACCAGAAGTCGGGAA  
 AGTCAAAGGGGAAGGATGTTAACTTGGCGGAGTTCGCTGTGGCTGCAGGGGACCAGATGCTT  
 TACAGGAGTGAGGACATCCAGCTAGATTACAAAAACAACATCCTGAAGGAGAGGGCGGAGCT  
 GGCCACAGCCCCCTGCCTGCCAAGTACATCGACCTAGACAAAGGGTTCCGGAAGGAGA  
 ACTGCAAA**TAG**GGAGGCCCTGGGCTCCTGGCTGGGCCAGCAGCTGCACCTCTCCTGTCTGTGCTC  
 CTCGGGGCATCTCCTGATGCTCCGGGGCTCACCCCCCTTCCAGCGGCTGGTCCCGCTTTCCT  
 GGAATTTGGCCTGGGCGTATGCAGAGGCCGCTCCACACCCCTCCCCAGGGGCTTGGTGGC  
 AGCATAGCCCCACCCCTGCGGCCCTTTGCTCACGGGTGGCCCTGCCACCCCTGGCACAACC  
 AAAATCCCACTGATGCCCATCATGCCCTCAGACCCTTCTGGGCTCTGCCCGCTGGGGGCTG  
 AAGACATTCCTGGAGGACACTCCCATCAGAACCTGGCAGCCCCAAAACCTGGGGTCAGCCTCA  
 GGGCAGGAGTCCCACTCCTCCAGGGCTCTGCTCGTCCGGGGCTGGGAGATGTTCTGGAGGA  
 GGACACTCCCATCAGAACTTGGCAGCCTTGAAGTTGGGGTCAGCCTCGGCAGGAGTCCCACT  
 CCTCCTGGGGTGCTGCCTGCCACCAAGAGCTCCCCACCTGTACCACCATGTGGGACTCCAG  
 GCACCATCTGTTCTCCCCAGGGACCTGCTGACTTGAATGCCAGCCCTTGCTCCTCTGTGTTG  
 CTTTGGGGCACCTGGGGCTGCACCCCCCTGCCCTTCTCTGCCCCATCCCTACCCTAGCCTTG  
 CTCTCAGCCACCTTGATAGTCACTGGGCTCCCTGTGACTTCTGACCCTGACACCCCTCCCTT  
 GGACTCTGCCTGGGCTGGAGTCTAGGGCTGGGGCTACATTTGGCTTCTGTACTGGCTGAGGA  
 CAGGGGAGGGAGTGAAGTTGGTTTGGGGTGGCCTGTGTTGCCACTCTCAGCACCCACATTT  
 GCATCTGCTGGTGGACCTGCCACCATCACAATAAAGTCCCCATCTGATTTTTAAAAA  
 AAAAA

**FIGURE 46**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60618

<subunit 1 of 1, 341 aa, 1 stop

<MW: 38070, pI: 6.88, NX(S/T): 1

MALPSRILLWKLVLQSSAVLLHSAVEETDAGLYTCNLHHHYCHLYESLAVRLEVTDGPPAT  
PAYWDGEKEVLAVARGAPALLTCVNRGHVWTDHRVVEEAQQVVHWRQPPGVPHDRADRLDL  
YASGERRAYGPLFLRDRVAVGADAFERGFSLRIEPLVADEGTYSCHLHHHYCGLHERRVF  
HLTVAEPHAEPPIRGSPGNGSSHSGAPGPDPTLARGHNVINVIVPESRAHFFQQLGYVLATL  
LLFILLLVTVLLAARRRRGGYEYSDQKSGKSKGKDVNLAEFAVAAGDQMLYRSEDIQLDYKN  
NILKERAELAHSPLPAKYIDLKGRKENCK

**Important features:****Signal peptide:**

amino acids 1-19

**Transmembrane domain:**

amino acids 237-262

**N-glycosylation site.**

amino acids 205-208

**Cell attachment sequence.**

amino acids 151-154

**Coproporphyrinogen III oxidase proteins.**

amino acids 115-140

**FIGURE 47**

CGCCGGAGGCAGCGGGCGGCGTGGCGCAGCGGCGACATGGCCGTTGTCTCAGAGGACGACTTT  
CAGCACAGTTCAAACCTCCACCTACGGAACCACAAGCAGCAGTCTCCGAGCTGACCAGGAGGC  
ACTGCTTGAGAAGCTGCTGGACCGCCCGCCCCCTGGCCTGCAGAGGCCCGAGGACCGCTTCT  
GTGGCACATACATCATCTTCTTCAGCCTGGGCATTGGCAGTCTACTGCCATGGAACTTCTTT  
ATCACTGCCAAGGAGTACTGGATGTTCAAACCTCCGCAACTCCTCCAGCCCAGCCACCGGGGA  
GGACCCTGAGGGCTCAGACATCCTGAACTACTTTGAGAGCTACCTTGCCGTTGCCTCCACCG  
TGCCCTCCATGCTGTGCCTGGTGGCCAACCTTCCTGCTTGTCAACAGGGTTGCAGTCCACATC  
CGTGTCTGGCCTCACTGACGGTCATCCTGGCCATCTTCATGGTGATAACTGCACTGGTGAA  
GGTGGACACTTCCTCCTGGACCCGTGGTTTTTTTTGCGGTCACCATTTGTCTGCATGGTGATCC  
TCAGCGGTGCCTCCACTGTCTTCAGCAGCAGCATCTACGGCATGACCGGCTCCTTTTCTATG  
AGGAACTCCCAAGCACTGATATCAGGAGGAGCCATGGGCGGGACGGTCAGCGCCGTGGCCTC  
ATTGGTGGACTTGGCTGCATCCAGTGATGTGAGGAACAGCGCCCTGGCCTTCTTCCTGACGG  
CCACCATCTTCCTCGTGCTCTGCATGGGACTCTACCTGCTGCTGTCCAGGCTGGAGTATGCC  
AGGTACTACATGAGGCCTGTTCTTGCGGCCCATGTGTTTTCTGGTGAAGAGGAGCTTCCCCA  
GGAATCCCTCAGTGCCCCCTTCGGTGGCCTCCAGATTCAATTGATTCCCACACACCCCCCTCTCC  
GCCCCATCCTGAAGAAGACGGCCAGCCTGGGCTTCTGTGTACCTACGTCTTCTTCATCACC  
AGCCTCATCTACCCCGCCGTCTGCACCAACATCGAGTCCCTCAACAAGGGCTCGGGCTCACT  
GTGGACCACCAAGTTTTTTCATCCCCCTCACTACCTTCCTCCTGTACAACTTTGCTGACCTAT  
GTGGCCGGCAGCTCACCGCCTGGATCCAGGTGCCAGGGCCCAACAGCAAGGCGCTCCCAGGG  
TTCGTGCTCCTCCGGACCTGCCTCATCCCCCTCTTCGTGCTCTGTAACCTACCAGCCCCGCGT  
CCACCTGAAGACTGTGGTCTTCAGTCCGATGTGTACCCCGCACTCCTCAGCTCCCTGCTGG  
GGCTCAGCAACGGCTACCTCAGCACCTGGCCCTCCTCTACGGGCCTAAGATTGTGCCCAGG  
GAGCTGGCTGAGGCCACGGGAGTGGTGATGTCTTTTATGTGTGCTTGGGCTTAACACTGGG  
CTCAGCCTGCTCTACCCTCCTGGTGCACCTCATCTAGAAGGGAGGACACAAGGACATTGGTG  
CTTCAGAGCCTTTGAAGATGAGAAGAGAGTGCAGGAGGGCTGGGGGCCATGGAGGAAAGGCC  
TAAAGTTTCACTTGGGGACAGAGAGCAGAGCACACTCGGGCCTCATCCCTCCCAAGATGCCA  
GTGAGCCACGTCCATGCCATTCCGTGCAAGGCAGATATTCCAGTCATATTAACAGAACT  
CCTGAGACAGTTGAAGAAGAAATAGCACAAATCAGGGGTACTCCCTTCACAGCTGATGGTTA  
ACATTCCACCTTCTTTCTAGCCCTTCAAAGATGCTGCCAGTGTTTCGCCCTAGAGTTATTACA  
AAGCCAGTGCCAAAACCCAGCCATGGGCTCTTTGCAACCTCCCAGCTGCGCTCATTCCAGCT  
GACAGCGAGATGCAAGCAAATGCTCAGCTCTCCTTACCCTGAAGGGGTCTCCCTGGAATGGA  
AGTCCCTGGCATGGTCAGTCTCAGGCCCAAGACTCAAGTGTGCACAGACCCCTGTGTTCT  
GCGGGTGAACAACTGCCCACTAACCAGACTGGAAAACCCAGAAAGATGGGCCTTCCATGAAT  
GCTTCATTCCAGAGGGACAGAGGGCCTCCCTGTGCAAGGGATCAAGCATGTCTGGCCTGGG  
TTTTCAAAAAAAGAGGGATCCTCATGACCTGGTGGTCTATGGCCTGGGTCAAGATGAGGGTC  
TTTCAGTGTTCTGTGTTTACAACATGTCAAAGCCATTGGTTCAAGGGCGTAATAAATACTTGC  
GTATTCAAAAA

**FIGURE 48**

MAVISEDDEQHSSNSTYGTSSSLRADQEALLEKLLDRPPPGQLQRPEDRFCGTYYIFFSLGI  
GSLLPWNFFITAKEYWMFKLRNSSSPATGEDPEGSDILNYFESYLAVASTVPSMLCLVANFL  
LVNRVAVHIRVLASLTVILAIIFMVITALVKVDTSSWTRGFFAVTIVCMVILSGASTVFSSSI  
YGMTGSFPMRNSQALISGGAMGGTVSAVASLVDLAASSDVNRNSALAFFLTATIFLVLCMGLY  
LLLSRLEYARYYMRPVLAHVFSGEEELPQDSLSAPSVASRFIDSHTPPLRPILKKTASLGF  
CVTYVFFITS LIYPAVCTNIESLNKSGSLWTTKFFIPLTTFLLYNFADLCGRQLTAWIQVP  
GPNSKALPGFVLLRTCLIPFLVLCNYQPRVHLKTVVFQSDVYPALLSSLLGLSNGYLSTLAL  
LYGPKIVPRELAEATGVVMSFYVCLGLTLGSACSTLLVHLI

**Transmembrane domain:**

amino acids 50-74 (type II), 105-127, 135-153, 163-183, 228-252,  
305-330, 448-472

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 49**

GACAGTGGAGGGCAGTGGAGAGGACCGCGCTGTCTTGCTGTCACCAAGAGCTGGAGACACCA  
 TCTCCCACCGAGAGTCA**ATG**GGCCCCATTGGCCCTGCACCTCCTCGTCCTCGTCCCCATCCTCC  
 TCAGCCTGGTGGCCTCCCAGGACTGGAAGGCTGAACGCAGCCAAGACCCCTTCGAGAAATGC  
 ATGCAGGATCCTGACTATGAGCAGCTGCTCAAGGTGGTGAACCTGGGGGCTCAATCGGACCCT  
 GAAGCCCCAGAGGGTGATTGTGGTTGGCGCTGGTGTGGCCGGGCTGGTGGCCGCCAAGGTGC  
 TCAGCGATGCTGGACACAAGGTACCATCCTGGAGGCAGATAACAGGATCGGGGGCCGCATC  
 TTCACCTACCGGGACCAGAACACGGGCTGGATTGGGGAGCTGGGAGCCATGCGCATGCCCAG  
 CTCTCACAGGATCCTCCACAAGCTCTGCCAGGGCCTGGGGCTCAACCTGACCAAGTTCACCC  
 AGTACGACAAGAACACGTGGACGGAGGTGCACGAAGTGAAGCTGCGCAACTATGTGGTGGAG  
 AAGGTGCCCCGAGAAGCTGGGCTACGCCTTGCGTCCCCAGGAAAAGGGCCACTCGCCCGAAGA  
 CATCTACCAGATGGCTCTCAACCAGGCCCTCAAAGACCTCAAGGCACTGGGCTGCAGAAAGG  
 CGATGAAGAAGTTTGAAAGGCACACGCTCTTGGAATATCTTCTCGGGGAGGGGAACCTGAGC  
 CGGCCGGCCGTGCAGCTTCTGGGAGACGTGATGTCCGAGGATGGCTTCTTCTATCTCAGCTT  
 CGCCGAGGCCCTCCGGGGCCACAGCTGCCTCAGCGACAGACTCCAGTACAGCCGCATCGTGG  
 GTGGCTGGGACCTGCTGCCGCGCGCGCTGCTGAGCTCGCTGTCCGGGCTTGTGCTGTTGAAC  
 GCGCCCGTGGTGGCGATGACCCAGGGACCGCACGATGTGCACGTGCAGATCGAGACCTCTCC  
 CCCGGCGCGGAATCTGAAGGTGCTGAAGGCCGACGTGGTGTGCTGACGGCGAGCGGACCGG  
 CGGTGAAGCGCATCACCTTCTCGCCCGCGCTGCCCCGCCACATGCAGGAGGCGCTGCGGAGG  
 CTGCACTACGTGCCGGCCACCAAGGTGTTCTTAAGCTTCCGCAGGCCCTTCTGGCGCGAGGA  
 GCACATTGAAGGCGGCCACTCAAACACCGATCGCCCGTCGCGCATGATTTTCTACCCGCCGC  
 CGCGCGAGGGCGCGCTGCTGCTGGCCTCGTACACGTGGTCGGACGCGGCGGCAGCGTTCCGCC  
 GGCTTGAGCCGGGAAGAGGCGTTGCGCTTGGCGCTCGACGACGTGGCGGCATTGCACGGGCC  
 TGTCGTGCGCCAGCTCTGGGACGGCACCGGCGTCGTCAAGCGTTGGGCGGAGGACCAGCACA  
 GCCAGGGTGGCTTTGTGGTACAGCCGCCGGCGCTCTGGCAAACCGAAAAGGATGACTGGACG  
 GTCCCTTATGGCCGCATCTACTTTGCCGGCGAGCACACCGCCTACCCGCACGGCTGGGTGGA  
 GACGGCGGTCAAGTCGGCGCTGCGCGCCGCCATCAAGATCAACAGCCGGAAGGGGCTGCAT  
 CGGACACGGCCAGCCCCGAGGGGCACGCATCTGACATGGAGGGGCAGGGGCATGTGCATGGG  
 GTGGCCAGCAGCCCCTCGCATGACCTGGCAAAGGAAGAAGGCAGCCACCCTCCAGTCCAAGG  
 CCAGTTATCTCTCCAAAACACGACCCACACGAGGACCTCGCAT**TAA**AGTATTTTCGGAAAAA  
 AA

**FIGURE 50**

MAPLALHLLVLPILLSLVLASQDWKAERSQDPFEKCMQDPDYEQLLKVVVTWGLNRTLKPQRV  
IVVGAGVAGLVAAKVLSDAGHKVTILEADNRIGGRIFTYRDQNTGWIGELGAMRMPSSHRIL  
HKLCQGLGLNLTKFTQYDKNTWTEVHEVKLRNYVVEKVPEKLGALRPQEKGHSPEDIYQMA  
LNQALKDLKALGCRKAMKKFERHTLLEYLLGEGNLSRPAVQLLGDVMSDGGFFYLSFAEALR  
AHSCLSDRQLQYSRIVGGWDLPRALLSSLSGLVLLNAPVVAMTQGPHDVHVQIETSPARNL  
KVLKADVLLLTASGPAVKRITFSPPLPRHMQEALRRLHYVPATKVFLSFRRPFWRREEHIEGG  
HSNTDRPSRMIFYPPPREGALLASYTWSDAAAAFAGLSREEALRLALDDVAALHGPVVRQL  
WDGTGVVKRWAEDQHSQGGFVVQPPALWQTEKDDWTVPYGRIYFAGEHTAYPHGWVETAVKS  
ALRAAIKINSRKGPASDTASPEGHASDMEGQGHVHGVASSPSHDLAKEEGSHPPVQGQLSLQ  
NTTHTRTSH

**Signal peptide:**

amino acids 1-21

**FIGURE 51**

CTGACATGGCCTGACTCGGGACAGCTCAGAGCAGGGCAGAACTGGGGACACTCTGGGCCGGCCTTCTGCCTGCAT  
 GGACGCTCTGAAGCCACCCTGTCTCTGGAGGAACACGAGCGAGGGAAGAAGGACAGGGACTCGTGTGGCAGGAA  
 GAACTCAGAGCCGGGAAGCCCCATTCACTAGAAGCACTGAGAGATGCGGCCCCCTCGCAGGGTCTGAATTTCT  
 GCTGCTGTTTCAAAAGATGCTTTTTATCTTTAACTTTTTGTTTTCCCCACTTCCGACCCCGCGTTGATCTGCAT  
 CCTGACATTTGGAGCTGCCATCTTCTTGTGGCTGATCACCAGACCTCAACCCGTCTTACCTCTTCTTGACCTGAA  
 CAATCAGTCTGTGGGAATTGAGGGAGGAGCACGGAAGGGGGTTTCCCAGAAGAACAATGACCTAACAAGTTGCTG  
 CTTCTCAGATGCCAAGACTATGTATGAGGTTTTCAAAGAGGACTCGCTGTGTCTGACAATGGGCCCTGCTTGGG  
 ATATAGAAAACCAACCAGCCCTACAGATGGCTATCTTACAAACAGGTGTCTGATAGAGCAGAGTACCTGGGTTT  
 CTGTCTCTTGCATAAAGGTTATAAATCATCACCAGACCAGTTTGTGCGCATCTTTGCTCAGAATAGGCCAGAGTG  
 GATCATCTCCGAATTGGCTTGTACACGTACTCTATGGTAGCTGTACCTCTGTATGACACCTTGGGACCAGAAGC  
 CATCGTACATATTGTCAACAAGGCTGATATCGCCATGGTGATCTGTGACACACCCCAAAAGGCATTGGTGCTGAT  
 AGGGAATGTAGAGAAAGGCTTCACCCCGAGCCTGAAGGTGATCATCCTTATGGACCCCTTTGATGATGACCTGAA  
 GCAAAGAGGGGAGAAGAGTGGAATTGAGATCTTATCCCTATATGATGCTGAGAACCTAGGCAAAGAGCACTTCAG  
 AAAACCTGTGCCTCCTAGCCCAGAAGACCTGAGCGTCATCTGCTTACCAGTGGGACCACAGGTGACCCCAAGG  
 AGCCATGATAACCCATCAAAATATTGTTTTCAAATGCTGCTGCCTTTCTCAAATGTGTGGAGCATGCTTATGAGCC  
 CACTCCTGATGATGTGGCCATATCCTACCTCCCTCTGGCTCATATGTTTGAGAGGATTGTACAGGCTGTTGTGTA  
 CAGCTGTGGAGCCAGAGTTGGATTCTTCCAAGGGGATATTGCGTTGCTGGCTGACGACATGAAGACTTTGAAGCC  
 CACATTGTTTTCCCGCGGTGCCTCGACTCCTTAACAGGATCTACGATAAGGTACAAAATGAGGCCAAGACACCCTT  
 GAAGAAGTTCTTGTGAAGCTGGCTGTTTCCAGTAAATTCAAAGAGCTTCAAAGGGTATCATCAGGCATGATAG  
 TTTCTGGGACAAGCTCATCTTTGCAAGATCCAGGACAGCCTGGGCGGAAGGGTTCTGTGAATTTGCTACTGGAGC  
 TGCCCCCATGTCCACTTCAGTCATGACATTCTTCCGGGCAGCAATGGGATGTGAGGTGTATGAAGCTTATGGTCA  
 AACAGAATGCACAGGTGGCTGTACATTTACATTACCTGGGGACTGGACATCAGGTACGTTGGGGTGCCCCCTGGC  
 TTGCAATTACGTGAAGCTGGAAGATGTGGCTGACATGAAGTACTTTACAGTGAATAATGAAGGAGAGGTCTGCAT  
 CAAGGGTACAAACGTGTTCAAAGGATACCTGAAGGACCCTGAGAAGACACAGGAAGCCCTGGACAGTGATGGCTG  
 GCTTCACACAGGAGACATTGGTCTGCTGGCTCCCGAATGGAACCTCTGAAGATCATCGACCGTAAAAAGAACATTTT  
 CAAGCTGGCCCAAGGAGAATACATTGCACCAGAGAAGATAGAAAATATCTACAACAGGAGTCAACCAGTGTTACA  
 AATTTTTGTACACGGGGAGAGCTTACGGTCATCCTTAGTAGGAGTGGTGGTTTCTGACACAGATGTACTTCCCTC  
 ATTTGTCAGCCAAGCTTGGGGTGAAGGGCTCCTTTGAGGAACTGTGCCAAAACCAAGTTGTAAGGGAAGCCATTTT  
 AGAAGACTTGCAGAAAATTTGGGAAAGAAAGTGGCCTTAAACTTTTGAACAGGTCAAAGCCATTTTCTTCATCC  
 AGAGCCATTTTCCATTGAAAATGGGCTCTTGACACCAACATTGAAAGCAAAGCGAGGAGAGCTTTCCAAATACTT  
 TCGGACCCAAATTGACAGCCTGTATGAGCACATCCAGGATTAGGATAAGGTACTTAAGTACCTGCCGGGCCACTG  
 TGCATGCTTGTGAGAAAATGGATTAAAAACTATTCTTACATTTGTTTTGCCTTTCCCTCCTATTTTTTTTTTAACC  
 TGTTAAACTCTAAAGCCATAGCTTTTTGTTTTATATTGAGACATATAATGTGTAAACTTAGTTCCTCAATATAATCA  
 ATCCTGTCTTTCCCATCTTCGATGTTGCTAATATTAAGGCTTCAGGGCTACTTTTATCAACATGCCTGTCTTCAA  
 GATCCCAGTTTATGTTCTGTGCTCCTTCCCTCATGATTTCCAACCTTAATACTATTAGTAACCACAAGTTCAAGGGT  
 CAAAGGGACCCCTCTGTGCCTTCTTCTTTGTTTTGTGATAAACATAACTTGCCAACAGTCTCTATGCTTATTTACA  
 TCTTCTACTGTTCAAATAAGAGATTTTTAAATTCTGAAAACTGCTTACAATTCATGTTTTCTAGCCACTCCAC  
 AAACCACTAAAATTTTAGTTTTAGCCTATCACTCATGTCAATCATATCTATGAGACAAATGTCTCCGATGCTCTT  
 CTGCGTAAATTAAATTGTGTACTGAAGGGAAAAGTTTGATCATACCAACATTTTCTTAAACTCTCTAGTTAGATA  
 TCTGACTTGGGAGTATTAAAAATTGGGTCTATGACATACTGTCCAAAAGGAATGCTGTTCTTAAAGCATTATTTA  
 CAGTAGGAACTGGGGAGTAAATCTGTTCCCTACAGTTTGTGCTGAGCTGGAAGCTGTGGGGGAAGGAGTTGACA  
 GGTGGGCCCAGTGAACCTTTCCAGTAAATGAAGCAAGCACTGAATAAAAACCTCCTGAACTGGGAACAAAGATCT  
 ACAGGCAAGCAAGATGCCACACAACAGGCTTATTTTCTGTGAAGGAACCAACTGATCTCCCCACCCTTGGATT  
 AGAGTTCCTGCTCTACCTTACCCACAGATAACACATGTTGTTTCTACTTGTAATGTAAAGTCTTTAAAATAAAC  
 TATTACAGATAAAAAA

**FIGURE 52**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60775

<subunit 1 of 1, 739 aa, 1 stop

<MW: 82263, pI: 7.55, NX(S/T): 3

MDALKPPCLWRNHERGKKDRDSCGRKNSEPGSPHSLEALRDAAPSQGLNFLLLFTKMLFI  
 FLFSPLPTPALICILTFGAAIFLWLITRPQPVLPPLDLNNQSVGIEGGARKGVSQKNNDLTS  
 CCFSDAKTMYEVFQRLAVSDNGPCLGYRKPNQPYRWLSYKQVSDRAEYLGSCLLHKGYKSS  
 PDQFVGIFAQNRPEWIISELACYTYSMVAVPLYDTLGPEAIVHIVNKADIAMVICDTPQKAL  
 VLIGNVEKGFTPSLKVIIIMDPFDDDLKQRGEKSGIEILSLYDAENLGKEHFRKPVPSPED  
 LSVICFTSGTTGDPKGAMITHQNIVSNAAFLKCVEHAYEPTPDDVAISYLPLAHMFERIVQ  
 AVVYSCGARVGGFFQGDIRLLADDMKTLKPTLFPVPRLLNRIYDKVQNEAKTPLKKFLLKLA  
 VSSKFKEQLQGIIRHDSFWDKLIFAKIQDSLGGRRVIVTGAAPMSTSVMTFFRAAMGCQVY  
 EAYGQTECTGGCTFTLPGDWTSGHVGVPACNYVKLEDVADMNYFTVNNEGEVCIKGTNVFK  
 GYLKDPEKTQEALDSDGWLHTGDIGRWLPNGTLKIIDRKKNIFKLAQGEYIAPEKIENIYNR  
 SQPVLQIFVHGESLRSSLVGVVVPDTPDLPSFAAKLGVKGSFEELCQNQVVRFAILEDLQKI  
 GKESGLKTFEQVKAIFLHPEPFSIENGLLTPTLKAKRGELSKYFRTQIDSLYEHIQD

**Important features:**

**Type II transmembrane domain:**

amino acids 61-80

**Putative AMP-binding domain signature.**

amino acids 314-325

**N-glycosylation site.**

amino acids 102-105, 588-591 and 619-622



**FIGURE 53**

T  
O  
T  
A  
L  
=  
6  
9  
3  
5  
4  
D  
D  
T

GGAGGCGGAGGCCGCGGCGAGCCGGGCCGAGCAGTGAGGGCCCTAGCGGGGCCCCGAGCGGGG  
 CCGGGGGCCCCCTAAGCCATTCTGAAGTCATGGGCTGGCCAGGACATTGGTGACCCGCCAAT  
 CCGGT**ATG**GACGACTGGAAGCCCAGCCCCCTCATCAAGCCCTTTGGGGCTCGGAAGAAGCGG  
 AGCTGGTACCTTACCTGGAAGTATAAACTGACAAACCAGCGGGGCCCTGCGGAGATTCTGTCA  
 GACAGGGGGCCGTGCTTTTCCTGCTGGTGACTGTCATTGTCAATATCAAGTTGATCCTGGACA  
 CTCGGCGAGCCATCAGTGAAGCCAATGAAGACCCAGAGCCAGAGCAAGACTATGATGAGGCC  
 CTAGGCCGCTGGAGCCCCCACGGCGCAGAGGCAGTGGTCCCCGGCGGGTCTGGACGTAGA  
 GGTGTATTCAAGTCGCAGCAAAGTATATGTGGCAGTGGATGGCACCACGGTGCTGGAGGATG  
 AGGCCCGGGAGCAGGGCCGGGGCATCCATGTCAATTGTCTCAACCAGGCCACGGGCCACGTG  
 ATGGCAAACGTGTGTTTGACACGTACTCACCTCATGAGGATGAGGCCATGGTGCTATTTCCT  
 CAACATGGTAGCGCCCGGCCGAGTGCTCATCTGCACTGTCAAGGATGAGGGCTCCTTCCACC  
 TCAAGGACACAGCCAAGGCTCTGCTGAGGAGCCTGGGCAGCCAGGCTGGCCCTGCCCTGGGC  
 TGGAGGGACACATGGGCCTTCGTGGGACGAAAAGGAGGTCTGTCTTCGGGGAGAAACATTC  
 TAAGTCACCTGCCCTCTCTTCCTGGGGGGACCCAGTCCTGCTGAAGACAGATGTGCCATTGA  
 GCTCAGCAGAAGAGGCAGAGTGCCACTGGGCAGACACAGAGCTGAACCGTCGCCCGCCGGCGC  
 TTCTGCAGCAAAGTTGAGGGCTATGGAAGTGTATGCAGCTGCAAGGACCCACACCCATCGA  
 GTTCAGCCCTGACCCACTCCCAGACAACAAGGTCCTCAATGTGCCTGTGGCTGTCAATTGCAG  
 GGAACCGACCCAATTACCTGTACAGGATGCTGCGCTCTCTGCTTTTCAGCCAGGGGGTGTCT  
 CCTCAGATGATAACAGTTTTTCATTGACGGCTACTATGAGGAACCCATGGATGTGGTGCCACT  
 GTTTGGTCTGAGGGGCATCCAGCATACTCCCATCAGCATCAAGAATGCCCGCGTGTCTCAGC  
 ACTACAAGGCCAGCCTCACTGCCACTTTCAACCTGTTTCCGGAGGCCAAGTTTGCTGTGGTT  
 CTGGAAGAGGACCTGGACATTGCTGTGGATTTTTTTCAGTTTCCTGAGCCAATCCATCCACCT  
 ACTGGAGGAGGATGACAGCCTGTACTGCATCTCTGCCTGGAATGACCAGGGGTATGAACACA  
 CGGCTGAGGACCCAGCACTACTGTACCGTGTGGAGACCATGCCTGGGCTGGGCTGGGTGCTC  
 AGGAGGTCCTTGTACAAGGAGGAGCTTGAGCCCAAGTGGCCTACACCGGAAAAGCTCTGGGA  
 TTGGGACATGTGGATGCGGATGCCTGAACAACGCCGGGGCCGAGAGTGCATCATCCCTGACG  
 TTTCCCGATCCTACCACTTTGGCATCGTCGGCCTCAACATGAATGGCTACTTTACAGAGGCC  
 TACTTCAAGAAGCACAAGTTCAACACGGTTCCAGGTGTCCAGCTCAGGAATGTGGACAGTCT  
 GAAGAAAGAAGCTTATGAAGTGGAAGTTCACAGGCTGCTCAGTGAGGCTGAGGTTCTGGACC  
 ACAGCAAGAACCCTTGTGAAGACTCTTTCCTGCCAGACACAGAGGGCCACACCTACGTGGCC  
 TTTATTTCGAATGGAGAAAGATGATGACTTCACCACCTGGACCCAGCTTGCCAAGTGCCCTCCA  
 TATCTGGGACCTGGATGTGCGTGGCAACCATCGGGGCCTGTGGAGATTGTTTCGGAAGAAGA  
 ACCACTTCCTGGTGGTGGGGGTCCCCGCTTCCCCCTACTCAGTGAAGAAGCCACCCTCAGTC  
 ACCCAATTTTCCTGGAGCCACCCCCAAAGGAGGAGGGAGCCCCAGGAGCCCCAGAACAGAC  
**ATGA**GACCTCCTCCAGGACCCTGCGGGGCTGGGTACTGTGTACCCCCAGGCTGGCTAGCCCT  
 TCCCTCCATCCTGTAGGATTTTGTAGATGCTGGTAGGGGCTGGGGCTACCTTGTTTTTAACA  
 TGAGACTTAATTACTAACTCCAAGGGGAGGGTTCCCTGCTCCAACACCCCGTTCTGAGTT  
 AAAAGTCTATTTATTTACTTCCTTGTGGAGAAGGGCAGGAGAGTACCTGGGAATCATTACG  
 ATCCCTAGCAGCTCATCCTGCCCTTTGAATACCTCACTTTCCAGGCCTGGCTCAGAATCTA  
 ACCTATTTATTGACTGTCCTGAGGGCCTTGAAAACAGGCCGAACCTGGAGGGCCTGGATTTT  
 TTTTGGGCTGGAATGCTGCCCTGAGGGTGGGGCTGGCTCTTACTCAGGAACTGCTGTGCC  
 CAACCCATGGACAGGCCAGCTGGGGCCCACATGCTGACACAGACTCACTCAGAGACCCTTA  
 GACACTGGACCAGGCCTCCTCTCAGCCTTCTCTTTGTCCAGATTTCCAAAGCTGGATAAGTT  
 GGTCATTGATTAAAAAAGGAGAAGCCCTCTGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 54**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA61185

><subunit 1 of 1, 660 aa, 1 stop

><MW: 75220, pI: 6.76, NX(S/T): 0

MDDWKPSPLIKPFGARKKRSWYLTWKYKLTNQALRRFCQTGAVLELLVTIVVNIKLILDTR  
RAISEANEDPEPEQDYDEALGRLEPPRRRGSGPRRVLDVEVYSSRSKVYVAVDGTTVLEDEA  
REQGRGIHVIVLNQATGHVMAKRVEDTYSPEDEAMVLFNLMVAPGRVLICTVKDEGSFHLK  
DTAKALLRSLGSQAGPALGWRDTWAFVGRKGGPVFGEKHSKSPALSSWGDPVLLKTDVPLSS  
AEEAECHWADTELNRNRRRRFCCKVEGYGSVCCKDPTPIEFSPDPLPDNKVLNVPVAVIAGN  
RPNYLYRMLRSLLSAQGVSPQMITVFIDGYEPMDDVVALFGLRGIQHTPISIKNARVSQHY  
KASLTATFNLFPFAKFAVVLEEDLDIAVDFFSFLSQSIHLLEEDDSLYCISAWNDQGYEHTA  
EDPALLYRVETMPGLGWVLRRLSYKEELEPKWPTPEKLWDWDMWMMRMPEQRRGRECIIPDVS  
RSYHFGIVGLNMNGYFHEAYFKKHKFNTVPGVQLRNVDLSLKKEAYEVEVHRLLEAEVLDHS  
KNPCEDSFLPDTEGHTYVAFIRMEKDDDFTTWTQLAKCLHIWDLDVIRGNHRGLWRLFRKKNH  
FLVVGVPASPYSVKKPPSVTPIFLEPPPKEEGAPGAPEQT

**Important features of the protein:**

**Transmembrane domain:**

amino acids 38-55

**Homologous region to Mouse GNT1**

amino acids 229-660

**FIGURE 55**

CGGACGCGTGGGCTGCTGGTGGGAAGGCCTAAAGAACTGGAAAGCCCACTCTCTTGGAACCACCACAC  
 CTGTTTAAAGAACCTAAGCACCATTAAAGCCACTGGAAATTTGTTGTCTAGTGGTTGTGGGTGAATA  
 AAGGAGGGCAGAAATGATGATTTTCATCTCCATTAGCCTGCTGTCTCTGGCTATGTTGGTGGGATGTTA  
 CGTGGCCGGAATCATTCCCTTGGCTGTTAATTTCTCAGAGGAACGACTGAAGCTGGTGAAGTGTGTTG  
 GTGCTGGCCTTCTCTGTGGAACCTGCTGTCAGTCTGCTGCCTGAAGGAGTACATGCCCTTTATGAA  
 GATATTCTTGAGGGAAAACACCACCAAGCAAGTGAACACATAATGTGATTGCATCAGACAAAGCAGC  
 AGAAAAATCAGTTGTCCATGAACATGAGCAGAGCCACGACCACACAGCTGCATGCCATATATTGGTG  
 TTTCCCTCGTTCTGGGCTTCGTTTTTCATGTTGCTGGTGGACCAGATTGGTAACTCCCATGTGCATTCT  
 ACTGACGATCCAGAAGCAGCAAGGTCTAGCAATTCAAAATCACCACCACGCTGGGTCTGGTTGTCCA  
 TGCTGCAGCTGATGGTGTGCTTTGGGAGCAGCAGCATCTACTTCACAGACCAGTGTCCAGTTAATTG  
 TGTTTGTGGCAATCATGCTACATAAGGCACCAGCTGCTTTTGGACTGGTTTCTTCTTGATGCATGCT  
 GGCTTAGAGCGGAATCGAATCAGAAAGCACTTGTGGTCTTTGCATTGGCAGCACCAGTTATGTCCAT  
 GGTGACATACTTAGGACTGAGTAAGAGCAGTAAAGAAGCCCTTTCAGAGGTGAACGCCACGGGAGTGG  
 CCATGCTTTTCTCTGCCGGGACATTTCTTTATGTTGCCACAGTACATGTCCTCCCTGAGGTGGGCGGA  
 ATAGGGCACAGCCACAAGCCCGATGCCACGGGAGGGAGAGGCCTCAGCCGCCTGGAAGTGGCAGCCCT  
 GGTTCTGGGTTGCCTCATCCCTCTCATCTGTGCTAGTAGGACACCAGCATTAATGTTCAAGGTCCAGC  
 CTTGGTCCAGGGCCGTTTGCCATCCAGTGAGAACAGCCGGCACGTGACAGCTACTCACTTCCCTCAGTC  
 TCTTGTCTCACCTTGCGCATCTCTACATGTATTCCCTAGAGTCCAGAGGGGAGGTGAGGTTAAACCTG  
 AGTAATGGAAAAGCTTTTAGAGTAGAAACACATTTACGTTGCAGTTAGCTATAGACATCCCATTGTGT  
 TATCCTTTTAAAGGCCCTTGACATTTTGCGTTTTAATATTTCTCTTAACCCTATTCTCAGGGAAGATG  
 GAATTTAGTTTTTAAGGAAAAGAGGAGAACCTCATACTCACAATGAAATAGTGATTATGAAAATACAGT  
 GTTCTGTAATTAAGCTATGTCTCTTTCTTCTTCTTAGGCTCTGCTACTTTATCCATTGATTTTTT  
 AACATGGTTCCACCATGTAAGACTGGTGTCTTTAGCATCTATGCCACATGCGTTGATGGAAGGTGATA  
 GCACCCACTCACTTAGATGCTAAAGGTGATTCTAGTTAATCTGGGATTAGGGTCAGGAAAATGATAGC  
 AAGACACATTGAAAGCTCTCTTTATACTCAAAAGAGATATCCATTGAAAAGGGATGTCTAGAGGGATT  
 TAAACAGCTCCTTTGGCACGTGCCTCTCTGAATCCAGCCTGCCATTCCATCAAATGGAGCAGGAGAGG  
 TGGGAGGAGCTTCTAAAGAGGTGACTGGTATTTTGTAGCATTCTTGTCAAGTTCTCCTTTGCAGAAT  
 ACCTGTCTCCACATTCCTAGAGAGGAGCCAAGTTCTAGTAGTTTCAGTTCTAGGCTTTCCTTCAAGAA  
 CAGTCAGATCACAAGTGTCTTTGGAAATTAAGGGATATTAAATTTTAAAGTGAATTTTGGATGGTTAT  
 TGATATCTTTGTAGTAGCTTTTTTTTAAAGACTACCAAATGTATGGTTGTCTTTTTTTTTTGT  
 TTTTTTTTTAATATTTCTCTTAGCAGATCAGCAATCCCTCTAGGGACCTAAATACTAGGTCAGCTTT  
 GCGACACTGTGTCTTCTCACATAACCACCTGTAGCAAGATGGATCATAAATGAGAAGTGTGTCCTA  
 TTGATTTTAAAGCTTATTGGAATCATGTCTCTTGTCTCTTCGTCTTTTCTTTGCTTTTCTTCTA  
 TCCCTCTAGCCTCTCCTCGCCACAATTTGCTGCTTACTGCTGGTGTAAATATTGTGTGGGATGAATT  
 CTTATCAGGACAACCACTTCTCGAACTGTAATAATGAAGATAATAATATCTTTATTCTTTATCCCTT  
 CAAAGAAATTACCTTTGTGTCAAATGCCGCTTTGTTGAGCCCTTAAAATACCACCTCCTCATGTGTAA  
 ATTGACACAATCACTAATCTGGTAATTTAAACAATTGAGATAGCAAAAGTGTTTAACAGACTAGGATA  
 ATTTTTTTTTTTCATATTTGCCAAAATTTTGTAAACCTGTCTTGTCAAATAAGTGTATAATATTGTAT  
 TATTAATTTATTTTACTTTCTATACCATTTCAAACACATTACACTAAGGGGGAACCAAGACTAGTT  
 TCTTCAGGGCAGTGGACGTAGTAGTTTGTAAACGTTTCTATGACGCATAAGCTAGCATGCCTATG  
 ATTTATTTCTTTCATGAATTTGTCACTGGATCAGCAGCTGTGGAAATAAAGCTTGTGAGCCCTCTGCT  
 GGCCACAGTGAGGAAAGTAGCACAAATAGGATACAGTTGTATGTAGTCATTGGCAACAATTGCATACA  
 ATTTTACTACCAAGAGAAGGTATAGTATGGAAAGTCAAATGACTTCTTGTATTGGATGTTAACAGCT  
 GACTGGTGTGAGACTTGAGGTTTCATCTAGTCCTTCAAACCTATATGGTTGCCTAGATTCTCTCTGGA  
 AACTGACTTTGTCAAATAAATAGCAGATTGTAGTGTCAAAAAAA

**FIGURE 56**

MDDFISISLLSLAMLVGCVAGIIPLAVNFSEERLKLVTVLGAGLLCGTALAVIVPEGVHAL  
YEDILEGKHHQASETHNVIASDKAAEKSVVHEHEHSHDHTQLHAYIGVSLVLGFVFMLLVDQ  
IGNSHVHSTDDPEAARSSNSKITTTLGLVVHAAADGVALGAAASTSQTTSVQLIVFVAIMLHK  
APAAFGLVSFLMHAGLERNRIRKHLVLFALAAPVMSMVTYLGLSKSSKEALSEVNATGVAML  
FSAGTFLYVATVHVLPEVGGIGHSHKPDATGGRGLSRLEVAALVLGCLIPLIILSVGHQH

**Signal peptide:**

amino acids 1-18

**Transmembrane domain:**

amino acids 37-56, 106-122, 211-230, 240-260, 288-304

**FIGURE 57**

GCTCGAGGCCGGCGGCGGCGGGAGAGCGACCCGGGCGGCCTCGTAGCGGGGCCCCGGATCCC  
 CGAGTGGCGGCCGGAGCCTCGAAAAGAGATTCTCAGCGCTGATTTTGAGATGATGGGCTTGG  
 GAAACGGGCGTCGCAGCATGAAGTCGCCGCCCTCGTGCTGGCCGCCCTGGTGGCCTGCATC  
 ATCGTCTTGGGCTTCAACTACTGGATTGCGAGCTCCCGGAGCGTGGACCTCCAGACACGGAT  
 CATGGAGCTGGAAGGCAGGGTCCGCAGGGCGGCTGCAGAGAGAGGGCGCCGTGGAGCTGAAGA  
 AGAACGAGTTCCAGGGAGAGCTGGAGAAGCAGCGGGAGCAGCTTGACAAAATCCAGTCCAGC  
 CACAACTTCCAGCTGGAGAGCGTCAACAAGCTGTACCAGGACGAAAAGGCGGTTTTTGGTGAA  
 TAACATCACCACAGGTGAGAGGCTCATCCGAGTGCTGCAAGACCAGTTAAAGACCCTGCAGA  
 GGAATTACGGCAGGCTGCAGCAGGATGTCCTCCAGTTTCAGAAGAACCAGACCAACCTGGAG  
 AGGAAGTTCTCCTACGACCTGAGCCAGTGCATCAATCAGATGAAGGAGGTGAAGGAACAGTG  
 TGAGGAGCGAATAGAAGAGGTCACCAAAAAGGGGAATGAAGCTGTAGCTTCCAGAGACCTGA  
 GTGAAAACAACGACCAGAGACAGCAGCTCCAAGCCCTCAGTGAGCCTCAGCCCAGGCTGCAG  
 GCAGCAGGCCTGCCACACACAGAGGTGCCACAAGGGAAGGGAACGTGCTTGGTAACAGCAA  
 GTCCCAGACACCAGCCCCCAGTTCCGAAGTGGTTTTTGGATTCAAAGAGACAAGTTGAGAAAG  
 AGGAAACCAATGAGATCCAGGTGGTGAATGAGGAGCCTCAGAGGGACAGGCTGCCGCAGGAG  
 CCAGGCCGGGAGCAGGTGGTGGAAAGACAGACCTGTAGGTGGAAGAGGCTTCGGGGGAGCCGG  
 AGAACTGGGCCAGACCCACAGGTGCAGGCTGCCCTGTCAGTGAGCCAGGAAAATCCAGAGA  
 TGGAGGGCCCTGAGCGAGACCAGCTTGTCATCCCCGACGGACAGGAGGAGGAGCAGGAAGCT  
 GCCGGGGAAGGGAGAAACCAGCAGAACTGAGAGGAGAAGATGACTACAACATGGATGAAAA  
 TGAAGCAGAATCTGAGACAGACAAGCAAGCAGCCCTGGCAGGGAATGACAGAAACATAGATG  
 TTTTAAATGTTGAAGATCAGAAAAGAGACACCATAAATTTACTTGATCAGCGTGAAAAGCGG  
 AATCATACACTCTGAATTGAACTGGAATCACATATTTACAACAGGGCCGAAGAGATGACTA  
 TAAATGTTTCATGAGGGACTGAATACTGAAAACGTGAAATGTACTAAATAAATGTACATCTGA

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

**FIGURE 58**

MMGLGNRRSMKSPPLVLAALVACIIIVLGFNYWIIASSRSVDLQTRIMELEGRVRRAAAERGA  
VELKKNEFQGELEKQREQLDKIQSSSHNFQLESVNKLYQDEKAVLVNNITTGERLIRVLQDQL  
KTLQARNYGRLQQDVLQFQKNQTNLERKFSYDLSQCINQMKEVKEQCEERIEEVTKKGNEAVA  
SRDLSENNDQRQQQLQALSEPQPRQLQAAGLPHTTEVPQGKGNVLGNSKSQTPAPSSEVVLDSCR  
QVEKEETNEIQVVNEEPQRDRLPQEPGREQVVEDRPVGGRGFGGAGELGQTPQVQAALSVSQ  
ENPEMEGPERDQLVIPDGQEEEQEAAGEGRNQQLRGEDDYNMDENEAESETDKQAALAGND  
RNIDVFENVEDQKRD TINLLDQREKRNHTL

**Signal peptide:**

amino acids 1-29

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 59**

GGATGCAGAAAGCCTCAGTGTGTGCTCTTCTGCGCTGGGTCTGCTTCTCTTCTACGCTGGCATTGCCCTCTTCA  
 CCAGTGGCTTCTGCTCACCCGTTTGGAGCTCACCACCATAGCAGCTGCCAAGAGCCCCAGGCCCTGGGTCCC  
 TGCCATGGGGGAGCCAAGGGAAACCTGGGGCCTGCTGGATGGCTTCCCGATTTTCGCGGGTTGTGTTGGTGCTGA  
 TAGATGCTCTGCGATTTGACTTCGCCCAGCCCCAGCATTACACGTGCCTAGAGAGCCTCCTGTCTCCCTACCCT  
 TCCTGGGCAAACCTAAGCTCCTTGACAGAGGATCCTGGAGATTACAGCCCCACCATGCCCCGCTCTACCGATCTCAGG  
 TTGACCCTCCTACCACCACCATGCAGCGCCTCAAGGCCCTCACCCTGGCTCACTGCCTACCTTTATTGATGCTG  
 GTAGTAACCTCGCCAGCCACGCCATAGTGAAGACAATCTCATTAAAGCAGCTCACCAGTGCAGGAAGGCGTGTAG  
 TCTTCATGGGAGATGATACCTGGAAGACCTTTTCCCTGGTGCTTTCTCCAAAGCTTTCTTCTTCCCATCCTTCA  
 ATGTCAGAGACCTAGACACAGTGGACAATGGCATCCTGGAACACCTCTACCCCACCATGGACAGTGGTGAATGGG  
 ACGTGCTGATTGCTCACTTCTTGGGTGTGGACCACTGTGGCCACAAGCATGGCCCTCACCACCCTGAAATGGCCA  
 AGAACTTAGCCAGATGGACCAGGTGATCCAGGGACTTGTGGAGCGTCTGGAGAATGACACACTGCTGGTAGTGG  
 CTGGGGACCATGGGATGACCACAAATGGAGACCATGGAGGGGACAGTGAGCTGGAGGTCTCAGCTGCTCTCTTTC  
 TGTATAGCCCCACAGCAGTCTTCCCCAGCACCCCCACAGAGGAGCCAGAGGTGATTCTCAAGTTAGCCTTGTGC  
 CCACGCTGGCCCTGCTGCTGGGCTGCCATCCATTTGGGAATATCGGGGAAGTGTGGCTGAGCTATTCTCAG  
 GGGGTGAGGACTCCCAGCCCCACTCCTCTGCTTTAGCCCCAAGCCTCAGCTCTCCATCTCAATGCTCAGCAGGTGT  
 CCCCATTTCTTCATACCTACTCAGCTGCTACTCAGGACCTTCAAGCTAAGGAGCTTCATCAGCTGCAGAACCTCT  
 TCTCCAAGGCCTCTGCTGACTACCACTGGCTTCTCCAGAGCCCCAAGGGGGCTGAGGCGACACTGCCGACTGTGA  
 TTGCTGAGCTGCAGCAGTTCCTGCGGGGAGCTCGGGCCATGTGCATCGAGTCTTGGGCTCGTTTCTCTCTGGTCC  
 GCATGGCGGGGGTACTGCTCTCTTGGCTGCTTCTGCTTTATCTGCCTGCTGGCATCTCAGTGGGCAATATCCC  
 CAGGCTTTCCATTCTGCCCTTACTCCTGACACCTGTGGCCTGGGGCCTGGTTGGGGCCATAGCGTATGCTGGAC  
 TCCTGGGAACTATTGAGCTGAAGCTAGATCTAGTGCTTCTAGGGGCTGTGGCTGCAGTGAGCTCATTCCCTCCCTT  
 TTCTGTGGAAGCCTGGGCTGGCTGGGGGTCCAAGAGGCCCTGGCAACCCTGTTTCCCATCCCTGGGCCCCGTC  
 TGTTACTCCTGCTGTTTCGCTTGGCTGTGTTCTTCTCTGATAGTTTTGTTGTAGCTGAGGCCAGGGCCACCCCT  
 TCCTTTTGGGCTCATTCATCCTGCTCCTGCTTGTCCAGCTTCACTGGGAGGGCCAGCTGCTTCCACCTAAGCTAC  
 TCACAATGCCCCGCTTGGCACTTCAGCCACAACAAACCCCCACGGCACAATGGTGCATATGCCCTGAGGCTTG  
 GAATTGGGTGCTTTTATGTACAAGGCTAGCTGGGCTTTTTTATCGTTGCCCTGAAGAGACACCTGTTTGGCACT  
 CCTCTCCCTGGCTGAGTCTCTGGCATCCATGGTGGGTGGTGCAGCCAGAATTTATGGTATGGAGCTTGTGTGG  
 CGGCGCTGGTGGCCCTGTAGCTGCCGTGCGCTTGTGGCTTCGCCGCTATGGTAATCTCAAGAGCCCCGAGCCAC  
 CCATGCTCTTGTGCGCTGGGGACTGCCCTAATGGCATTGGGTACTGCTGCCTACTGGGCATTGGCGTCGGGGG  
 CAGATGAGGCTCCCCCGCTCTCGGGCTCCTGGTCTCTGGGGCATCCATGGTGTGCTCGGGCTGTAGCAGGGC  
 TGGCTGCTTCAGGGCTCGCGCTGCTGCTTGAAGCCTGTGACAGTGTGGTGAAGGCTGGGGCAGGCGCTCCAA  
 GGACCAGGACTGTCTCACTCCCTTCTCAGGCCCCCCACTTCTCAAGCTGACTTGGATTATGTGGTCCCTCAAA  
 TCTACCGACACATGCAGGAGGAGTTCCGGGGCCGGTTAGAGAGGACCAAAATCTCAGGGTCCCCTGACTGTGGCTG  
 CTTATCAGTTGGGGAGTGTCTACTCAGCTGCTATGGTCACAGCCCTCACCCCTGTTGGCCTTCCCCTTCTGCTGT  
 TGCATGCGGAGCGCATCAGCCTTGTGTTCTGCTTCTGTTTCTGCAGAGCTTCTTCTCCTACATCTGCTTGTGCTG  
 CTGGGATACCCGTCACCACCCCTGGTCTTTTACTGTGCCATGGCAGGCAGTCTCGGCTTGGGCCCTCATGGCCA  
 CACAGACCTTCTACTCCACAGGCCACAGCCTGTCTTCCAGCCATCCATTGGCATGCAGCCTTCGTGGGATTCC  
 CAGAGGGTCATGGCTCCTGTACTTGGCTGCCTGCTTTGCTAGTGGGAGCCAACACCTTTGCCTCCCACCTCCTCT  
 TTGCAGTAGGTTGCCACTGCTCCTGCTCTGGCCTTTTCTGTGTGAGAGTCAAGGGCTGCGGAAGAGACAGCAGC  
 CCCCAGGGAATGAAGCTGATGCCAGAGTCAGACCCGAGGAGGAAGAGGAGCCACTGATGGAGATGCGGCTCCGGG  
 ATGCGCCTCAGCACTTCTATGCAGCACTGCTGCAGCTGGGCCTCAAGTACCTCTTTATCCTTGGTATTCAGATTC  
 TGGCCTGTGCCTTGGCAGCCTCCATCCTTCGAGGCATCTCATGGTCTGGAAAGTGTGTTGCCCTAAGTTCATAT  
 TTGAGGCTGTGGGCTTCATTGTGAGCAGCGTGGGACTTCTCCTGGGCATAGCTTTGGTGATGAGAGTGGATGGTG  
 CTGTGAGCTCCTGGTTCAGGCAGCTATTTCTGGCCCCAGCAGAGGTAGCCTAGTCTGTGATTACTGGCACTTGGCT  
 ACAGAGAGTGTGGAGAACAGTGTAGCCTGGCCTGTACAGGTACTGGATGATCTGCAAGACAGGCTCAGCCATAC  
 TCTTACTATCATGCAGCCAGGGGCCGCTGACATCTAGGACTTCATTATTCTATAATTAGGACCACAGTGGAGTA  
 TGATCCCTAACTCCTGATTTGGATGCATCTGAGGGACAAGGGGGGCGGTCTCCGAAGTGAATAAAATAGGCCGG  
 GCGTGGTGACTTGCACCTATAATCCCAGCACTTTGGGAGGCAGAGGTGGGAGGATTGCTTGGTCCCAGGAGTTCA  
 AGACCAGCCTGTGGAACATAACAAGACCCCGTCTCTACTATTTAAAAAAGTGAATAAAATGATAATAT

**FIGURE 60**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62809

<subunit 1 of 1, 1089 aa, 1 stop

<MW: 118699, pI: 8.49, NX(S/T): 2

MQKASVLLFLAWVCFLFYAGIALFTSGFLLTRLELTNHSSCQEPPGPGSLPWGSQGKPGACW  
 MASRFSRVVLVLIDALRFDFQAQPQSHVPREPPVSLPFLGKLSSLQRILEIQPHHARLYRSQ  
 VDPPTTTMQRLKALTGSLPTFIDAGSNFASHAIVEDNLIKQLTSAGRRVVFMGDDTWKDLF  
 PGAFSKAFFFPFNFVRDLDTVDNGILEHLYPTMDSGEWDVLIHFLGVDHCHGKHGPHHP  
 AKKLSQMDQVIQGLVERLENDTLLVVAGDHGTTNGDHGGDSELEVSAALFLYSPTAVFPST  
 PPEEPEVIPQVSLVPTLALLLGLPIPFNGNIGEVMAELFSGGEDSQPHSSALAQASALHLNAQ  
 QVSRLFHTYSAATQDLQAKELHQLQNLFSKASADYQWLLQSPKGAEATLPTVIAELQQFLRG  
 ARAMCIESWARFSLVRMAGGTALLAASCFICLLASQWAI SPGFPCPLLLTPVAWGLVGAIA  
 YAGLLGTIELKLDLVLLGAVAAVSSFLPFLWKAWAGWGSKRPLATLFPPIPGPVLLLLLLFRLA  
 VFFSDSFVVAEARATPFLGSGFILLLVVQLHWEGQLLPKLLTMPRLGTSATTNPPRHNGAY  
 ALRLGIGLLLCTRLAGLFHRCPEETPVCHSSPWLSPLASMVGGRAKNLWYGACVAALVALLA  
 AVRLWLRRYGNLKSPEPPMLFVRWGLPLMALGTAAYWALASGADEAPPRLRVLVSGASMLP  
 RAVAGLAASGLALLLWKPVTVLVKAGAGAPRTRTVLTPFSGPPTSQADLDYVVPQIYRHMQE  
 EFRGRLETKSQGPLTVAAYQLGSVYSAAMVTALTLLAFPLLLLHAERISLVFLLLFLQSFL  
 LLHLLAAGIPVTTPGPFTVPWQAVSAWALMATQTFYSTGHQPVFPAIHWHAAFVGFPEGHGS  
 CTWLPALLVGANTFASHLLFAVGCPLLLLWPFLCESQGLRKRQQPPGNEADARVRPEEEEEEP  
 LMEMRLRDAPQH FYAALLQLGLKYLFI LGIQILACALAASILRRHLMVWKVFAPKFIFEAVG  
 FIVSSVGLLLGIALVMRVDGAVSSWFRQLFLAQQR

**Important features:****Signal peptide:**

amino acids 1-16

**Transmembrane domains:**

amino acids 317-341, 451-470, 481-500, 510-527, 538-555, 831-850,  
 1016-1034, 1052-1070

**Leucine zipper pattern.**

amino acids 843-864

**N-glycosylation sites.**

amino acids 37-40, 268-271



## FIGURE 61

TGCGCTGCCGCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCTGT  
GTCTCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGT  
CCTACAATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATT  
TCATCACAAATTGGCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGT  
CCTGACAGCTCCAGAGAAGTGGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAACAAA  
TATACTCCAATCTGAAGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCC  
CAGTGTGTGACCAACCACACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGT  
ACACGTGGAGTCCTTCGTCCCAGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTG  
CCAGGACTTTGAAAGATCAATCATCAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTG  
CCCATATCTATTACCGTGTTTTCTTTTTTCTGTGATGGGCTATTCCATCTACCGATATATCCA  
CGTTGGCAAAGAGAAACACCCAGCAAATTTGATTTTGATTTATGGAAATGAATTTGACAAAA  
GATTCCTTTGTGCTGCTGAAAAAATCGTGATTAACCTTTATCACCTCAATATCTCGGATGAT  
TCTAAAATTTCTCATCAGGATATGAGTTTACTGGGAAAAAGCAGTGATGTATCCAGCCTTAA  
TGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGAGGAAGAGGAGGTGAAACATTTAG  
GGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAGAAAACACGGAAGGTACTTCT  
CTCACCCAGCAAGAGTCCCTCAGCAGAACAATACCCCCGATAAAACAGTCATTGAATATGA  
ATATGATGTCAGAACCACTGACATTTGTGCGGGGCCTGAAGAGCAGGAGCTCAGTTTGCAGG  
AGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAGTCTTGGGCCCCG  
CAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCTGGCGCAGGAGCA  
CACAGACTCGGAGGAGGGGCCGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGATCCCC  
AAACTGGCAGGCTGTGTATTCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCGAG  
CCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGC  
TCCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGT  
TATATGTGCAGATGGAAAACTGATGCCAACACTTCCTTTTGCCTTTTGTTCCTGTGCAAC  
AAGTGAGTCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTTCCAGT  
TTGTCAGTGTCTGTGAGAATTACTTATTTCTTTTCTCTATTCTCATAGCACGTGTGTGATTG  
GTTTCATGCATGTAGGTCTCTTAACAATGATGGTGGGCCTCTGGAGTCCAGGGGCTGGCCGGT  
TGTTCTATGCAGAGAAAGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTCAGG  
TGGGTGT

**FIGURE 62**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62815

<subunit 1 of 1, 442 aa, 1 stop

<MW: 49932, pI: 4.55, NX(S/T): 5

MSYNGLHQRVFKELKLLTLCSSISQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQ  
QIYSNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQ  
CARTLKDQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNEDF  
KRFFVPAEKIVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKH  
LGYASHLMEIFCDSEENTEGTSLTQQESLSRTIPDPKTVIEYDYDVRTTDICAGPEEQELSL  
QEEVSTQGTLLESQAALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWD  
PQTGRLCIPSLSSFDQDSEGCEPSEGDLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEW  
GLYVQMEN

**Important features:****Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 140-163

**N-glycosylation sites.**

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

**FIGURE 63**

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTCTCTGCGGGGAGACGCCAGCCTGCG  
 TCTGCC**ATG**GGGCTCGGGTTGAGGGGCTGGGGACGTCCTCTGCTGACTGTGGCCACCGCCCT  
 GATGCTGCCCCGTGAAGCCCCCGCAGGCTCCTGGGGGGCCAGATCATCGGGGGCCACGAGG  
 TGACCCCCCACTCCAGGCCCTACATGGCATCCGTGCGCTTCGGGGGGCCAACATCACTGCGGA  
 GGCTTCCTGCTGCGAGCCCGCTGGGTGGTCTCGGCCGCCCACTGCTTCAGCCACAGAGACCT  
 CCGCACTGGCCTGGTGGTGTGGGCGCCACGTCCTGAGTACTGCGGAGCCCACCCAGCAGG  
 TGTTTGGCATCGATGCTCTCACCACGCACCCCGACTACCACCCCATGACCCACGCCAACGAC  
 ATCTGCCTGCTGCGGCTGAACGGCTCTGCTGTCCTGGGCCCTGCAGTGGGGCTGCTGAGGCT  
 GCCAGGGAGAAGGGCCAGGCCCCCCACAGCGGGGACACGGTGCCGGGTGGCTGGCTGGGGCT  
 TCGTGTCTGACTTTGAGGAGCTGCCGCCTGGACTGATGGAGGCCAAGGTCCGAGTGCTGGAC  
 CCGGACGTCTGCAACAGCTCCTGGAAGGGCCACCTGACACTTACCATGCTCTGCACCCGCAG  
 TGGGGACAGCCACAGACGGGGCTTCTGCTCGGCCGACTCCGGAGGGCCCCCTGGTGTGCAGGA  
 ACCGGGCTCACGGCCTCGTTTCCTTCTCGGGCCTCTGGTGCGGCGACCCCAAGACCCCCGAC  
 GTGTACACGCAGGTGTCCGCCTTTGTGGCCTGGATCTGGGACGTGGTTTCGGCGGAGCAGTCC  
 CCAGCCCGGCCCCCTGCCTGGGACCACCAGGCCCCCAGGAGAAGCCGCC**TGA**GCCACAACCT  
 TGCGGCATGCAAATGAGATGGCCGCTCCAGGCCTGGAATGTTCCGTGGCTGGGCCCCACGGG  
 AAGCCTGATGTTCAAGGTTGGGGTGGGACGGGCAGCGGTGGGGCACACCCATTCCACATGCA  
 AAGGGCAGAAGCAAACCCAGTAAAATGTAACTGACAAAAAAAAAAAAAAAAAAAAAGAAA

## FIGURE 64

MGLGLRGWGRPLLTVATALMLPVKPPAGSWGAIIGGHEVTPHSRPYMASVRFGGQHHCGRF  
LLRARWVVSAAHCFSHRDLRTGLVVLGAHVLSTAEPTQQVFGIDALTTHPDYHPMTHANDIC  
LLRLNGSAVLGPAVGLLRPLPGRRARPPPTAGTRCRVAGWGFVSDFEELPPGLMEAKVRVLDPD  
VCNSSWKGHLTLTMLCTRSGDSHRRGFC SADSGGPLVCRNRAHGLVSFSGLWCGDPKTPDVY  
TQVSAFVAWIWDVVRRSSPQPGPLPGTTRPPGEAA

amino acids 1-30

**FIGURE 65**

GAGCTACCCAGGCGGCTGGTGTGCAGCAAGCTCCGCGCCGACTCCGGACGCCTGACGCCTGA  
 CGCCTGTCCCCGGCCCGGC**ATG**AGCCGCTACCTGCTGCCGCTGTCGGCGCTGGGCACGGTAG  
 CAGGCGCCCGCGTGTGCTCAAGGACTATGTACCCGGTGGGGCTTGCCCCAGCAAGGCCACC  
 ATCCCTGGGAAGACGGTCATCGTGACGGGGCGCCAACACAGGCATCGGGAAGCAGACCGCCTT  
 GGAAGTGGCCAGGAGAGGAGGCAACATCATCCTGGCCTGCCGAGACATGGAGAAGTGTGAGG  
 CGGCAGCAAAGGACATCCGCGGGGAGACCCTCAATCACCATGTCAACGCCCCGGCACCTGGAC  
 TTGGCTTCCCTCAAGTCTATCCGAGAGTTTGCAGCAAAGATCATTGAAGAGGAGGAGCGAGT  
 GGACATTCTAATCAACAACGCGGGTGTGATGCGGTGCCCCACTGGACCACCGAGGACGGCT  
 TCGAGATGCAGTTTGGCGTTAACCACCTGGGTCACTTTCTCTTGACAAACTTGCTGCTGGAC  
 AAGCTGAAAGCCTCAGCCCCCTTCGCGGATCATCAACCTCTCGTCCCTGGCCCATGTTGCTGG  
 GCACATAGACTTTGACGACTTGAAGTGGCAGACGAGGAAGTATAACACCAAAGCCGCCTACT  
 GCCAGAGCAAGCTCGCCATCGTCCTCTTCACCAAGGAGCTGAGCCGGCGGCTGCAAGGCTCT  
 GGTGTGACTGTCAACGCCCTGCACCCCGGCGTGGCCAGGACAGAGCTGGGCAGACACACGGG  
 CATCCATGGCTCCACCTTCTCCAGCACCACTCGGGCCCATCTTCTGGCTGCTGGTCAAGA  
 GCCCCGAGCTGGCCGCCAGCCCAGCACATACTGGCCGTGGCGGAGGAAGTGGCGGATGTT  
 TCCGGAAAGTACTTCGATGGACTCAAACAGAAGGCCCGGCCCGAGGCTGAGGATGAGGA  
 GGTGGCCCGGAGGCTTTGGGCTGAAAGTGCCCGCCTGGTGGGCTTAGAGGCTCCCTCTGTGA  
 GGGAGCAGCCCCCTCCCCAGAT**TAA**CCTCTGGAGCAGATTTGAAAGCCAGGATGGCGCCTCCAG  
 ACCGAGGACAGCTGTCCGCCATGCCCGCAGCTTCCCTGGCACTACCTGAGCCGGGAGACCCAG  
 GACTGGCGGCCGCCATGCCCGCAGTAGGTTCTAGGGGGCGGTGCTGGCCGCAGTGGACTGGC  
 CTGCAGGTGAGCACTGCCCCGGGCTCTGGCTGGTTCCGTCTGCTCTGCTGCCAGCAGGGGAG  
 AGGGGCCATCTGATGCTTCCCCTGGGAATCTAAACTGGGAATGGCCGAGGAGGAAGGGGCTC  
 TGTGCACTTGCAAGGCCACGTGAGGAGAGCCAGCGGTGCCTGTGCGGGAGGGTTCCAAGGTGC  
 TCCGTGAAGAGCATGGGCAAGTTGTCTGACACTTGGTGGATTCTTGGGTCCCTGTGGGACCT  
 TGTGCATGCATGGTCCCTCTCTGAGCCTTGGTTTCTTCAGCAGTGAGATGCTCAGAATAACTG  
 CTGTCTCCCATGATGGTGTGGTACAGCGAGCTGTTGTCTGGCTATGGCATGGCTGTGCCGGG  
 GGTGTTTGCTGAGGGGCTTCCCTGTGCCAGAGCCCAGCCAGAGAGCAGGTGCAGGTGTCATCCC  
 GAGTTCAGGCTCTGCACGGCATGGAGTGGGAACCCACAGCTGCTGCTACAGGACCTGGGA  
 TTGCCTGGGACTCCCACCTTCCCTATCAATTCTCATGGTAGTCCAAACTGCAGACTCTCAAC  
 TTGCTCATTT

1000  
 900  
 800  
 700  
 600  
 500  
 400  
 300  
 200  
 100  
 0

**FIGURE 66**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64842

><subunit 1 of 1, 331 aa, 1 stop

><MW: 35932, pI: 8.45, NX(S/T): 1

MSRYLLPLSALGTVAGAAVLLKDYVTGGACPSKATIPGKTVIVTGANTGIGKQTALELARRG  
GNIILACRDMEKCEAAAKDIRGETLNHHVNARHLDLASLKSIREFAAKIEEEEERVDILINN  
AGVMRCPHWTTEDGFEMQFGVNHLGHFLLTNLLLDKLGKASAPSRIINLSSLAHVAGHIDFDD  
LNWQTRKYNTKAAAYCQSKLAIVLFTKELSRRLQGSGVTVNALHPGVARTELGRHTGIHGSTF  
SSTTLGPIFWLLVKSPELAAQPSTYLAVAEELADVSGKYFDGLKQKAPAEAEDEEVARRLW  
AESARLVGLEAPSVREQLPR

**Signal peptide:**

amino acids 1-17

66/249

**FIGURE 67**

GAAGTTCGCGAGCGCTGGC**ATG**TGGTCCTGGGGCGCGGCTGGCGGCGCTGCTGGCGGTGCTG  
 GCGCTCGGGACAGGAGACCCAGAAAGGGCTGCGGCTCGGGGCGACACGTTCTCGGCGCTGAC  
 CAGCGTGGCGCGCGCCCTGGCGCCCCGAGCGCCGGCTGCTGGGGCTGCTGAGGCGGTACCTGC  
 GCGGGGAGGAGGCGCGGCTGCGGGACCTGACTAGATTCTACGACAAGGTACTTTCTTTGCAT  
 GAGGATTCAACAACCCCTGTGGCTAACCCCTCTGCTTGCAATTTACTCTCATCAAACGCCTGCA  
 GTCTGACTGGAGGAATGTGGTACATAGTCTGGAGGCCAGTGAGAACATCCGAGCTCTGAAGG  
 ATGGCTATGAGAAGGTGGAGCAAGACCTTCCAGCCTTTGAGGACCTTGAGGGAGCAGCAAGG  
 GCCCTGATGCGGCTGCAGGACGTGTACATGCTCAATGTGAAAGGCCTGGCCCCGAGGTGTCTT  
 TCAGAGAGTCACTGGCTCTGCCATCACTGACCTGTACAGCCCCAAACGGCTCTTTTCTCTCA  
 CAGGGGATGACTGCTTCCAAGTTGGCAAGGTGGCCTATGACATGGGGGATTATTACCATGCC  
 ATTCATGGCTGGAGGAGGCTGTCACTCTTCCGAGGATCTTACGGAGAGTGGAAGACAGA  
 GGATGAGGCAAGTCTAGAAGATGCCTTGGATCACTTGGCCTTTGCTTATTTCCGGGCAGGAA  
 ATGTTTCGTGTGCCCTCAGCCTCTCTCGGGAGTTTCTTCTCTACAGCCCAGATAATAAGAGG  
 ATGGCCAGGAATGTCTTGAAATATGAAAGGCTCTTGGCAGAGAGCCCCAACACAGTGGTAGC  
 TGAGGCTGTCATCCAGAGGCCCAATATACCCACCTGCAGACCAGAGACACCTACGAGGGGC  
 TATGTCAGACCTGGGTTCACAGCCCACTCTTACCAGATCCCTAGCCTCTACTGTTCTTAT  
 GAGACCAATTCCAACGCCTACCTGCTGCTCCAGCCCATCCGGAAGGAGGTCATCCACCTGGA  
 GCCCTACATTGCTCTCTACCATGACTTCGTCACTGACTCAGAGGCTCAGAAAATTAGAGAAC  
 TTGCAGAACCATGGCTACAGAGGTCACTGGTGGCATCAGGGGAGAAGCAGTTACAAGTGGAG  
 TACCGCATCAGCAAAAGTGCCTGGCTGAAGGACACTGTTGACCCAAAACCTGGTGACCCCTCAA  
 CCACCGCATTGCTGCCCTCACAGGCCTTGATGTCCGGCCTCCCTATGCAGAGTATCTGCAGG  
 TGGTGAACATATGGCATCGGAGGACACTATGAGCCTCACTTTGACCATGCTACGTCACCAAGC  
 AGCCCCCTCTACAGAATGAAGTCAGGAAACCGAGTTGCAACATTTATGATCTATCTGAGCTC  
 GGTGGAAGCTGGAGGAGCCACAGCCTTCATCTATGCCAACCTCAGCGTGCCTGTGGTTAGGA  
 ATGCAGCACTGTTTTGGTGGAACCTGCACAGGAGTGGTGAAGGGGACAGTGACACACTTCAT  
 GCTGGCTGTCCTGTCCTGGTGGGAGATAAGTGGGTGGCCAACAAGTGGATACATGAGTATGG  
 ACAGGAATTCCGCAGACCCCTGCAGCTCCAGCCCTGAAGACT**TGA**ACTGTTGGCAGAGAGAAGC  
 TGGTGGAGTCCCTGTGGCTTTCCAGAGAAGCCAGGAGCCAAAAGCTGGGGTAGGAGAGGAGAA  
 AGCAGAGCAGCCTCCTGGAAGAAGGCCTTGTCAGCTTTGTCTGTGCCTCGCAAATCAGAGGC  
 AAGGGAGAGGTTGTTACCAGGGGACACTGAGAATGTACATTTGATCTGCCCCAGCCACGGAA  
 GTCAGAGTAGGATGCACAGTACAAAGGAGGGGGGAGTGGAGGCCTGAGAGGGAAGTTTCTGG  
 AGTTCAGATACTCTCTGTTGGGAACAGGACATCTCAACAGTCTCAGGTTTCGATCAGTGGGTC  
 TTTTGGCACTTTGAACCTTGACCACAGGGACCAAGAAGTGGCAATGAGGACACCTGCAGGAG  
 GGGCTAGCCTGACTCCCAGAACTTTAAGACTTTCTCCCCACTGCCTTCTGCTGCAGCCCAAG  
 CAGGGAGTGTCCCCCTCCCAGAAGCATATCCCAGATGAGTGGTACATTATATAAGGATTTTT  
 TTTAAGTTGAAAACAACCTTTCTTTTCTTTTGTATGATGGTTTTTTAACACAGTCATTAAAA  
 ATGTTTATAAATCAAAA

**FIGURE 68**

MGPGARLAALLAVLALGTGDPERAAARGDTFSALTSVARALAPERLLGLLRRYLRGEEARL  
RDLTRFYDKVLSLHEDSTTPVANPLLAFTLIKRLQSDWRNVVHSLEASENIRALKDGYEKVE  
QDLPAFEDLEGAARALMRLQDVYMLNVKGLARGVFQRTGSAITDLYSPKRLFSLTGDDCFQ  
VGKVAYDMGDYYHAIPWLEEAVSLFRGSYGEWKTEDEASLEDALDHLAFAYFRAGNVSCALS  
LSREFLLYSPDNKRMARNVLKYERLLAESPNHVVAEAVIQRPNI PHLQTRD TYEGLCQTLGS  
QPTLYQIPSLYCSYETNSNAYLLLQPIRKEVIHLEPYIALYHDFVSDSEAQKIRELAEPWLQ  
RSVVASGEKQLQVEYRISKSAWLKDTVDPKLVTLNHRIAALTGLDVRPPYAEYLQVVNYGIG  
GHYEPHFDHATSPSSPLYRMKSGNRVATFMIYLSSVEAGGATAFIYANLSVPVVRNAALFWW  
NLHRSGEGDSDTLHAGCPVLVGDKWVANKWIHEYGQEFRRPCSSSPED

**Signal peptide:**

amino acids 1-19

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



**FIGURE 69**

GAGATAGGGAGTCTGGGTTTAAAGTTCCTGCTCCATCTCAGGAGCCCCTGCTCCCACCCCTAG  
 GAAGCCACCAGACTCCACGGTGTGGGGCCAATCAGGTGGAATCGGCCCTGGCAGGTGGGGCC  
 ACGAGCGCTGGCTGAGGGACCGAGCCGGAGAGCCCCGGAGCCCCCGTAACCCGCGCGGGGAG  
 CGCCCAGG**ATG**CCGCGCGGGGACTCGGAGCAGGTGCGCTACTGCGCGCGCTTCTCCTACCTC  
 TGGCTCAAGTTTTCACTTATCATCTATTCCACCGTGTCTGGCTGATTGGGGCCCTGGTCCT  
 GTCTGTGGGCATCTATGCAGAGGTTGAGCGGCAGAAATATAAAACCCCTTGAAAGTGCCTTCC  
 TGGCTCCAGCCATCATCCTCATCCTCCTGGGCGTCGTCATGTTTCATGGTCTCCTTCATTGGT  
 GTGCTGGCGTCCCTCCGTGACAACCTGTACCTTCTCCAAGCATTTCATGTACATCCTTGGGAT  
 CTGCCTCATCATGGAGCTCATTGGTGGCGTGGTGGCCTTGACCTTCCGGAACCAGACCATTG  
 ACTTCCTGAACGACAACATTCGAAGAGGAATTGAGAACTACTATGATGATCTGGACTTCAAA  
 AACATCATGGACTTTGTTCAGAAAAAGTTCAAGTGCTGTGGCGGGGAGGACTACCGAGATTG  
 GAGCAAGAATCAGTACCACGACTGCAGTGCCCCCTGGACCCCTGGCCTGTGGGGTGCCCTACA  
 CCTGCTGCATCAGGAACACGACAGAAAGTTGTCAACACCATGTGTGGCTACAAAACCTATCGAC  
 AAGGAGCGTTTTCAGTGTGCAGGATGTCATCTACGTGCGGGGCTGCACCAACGCCGTGATCAT  
 CTGGTTCATGGACAACCTACACCATCATGGCGTGCATCCTCCTGGGCATCCTGCTTCCCCAGT  
 TCCTGGGGGGTGCTGCTGACGCTGCTGTACATCACCCGGGTGGAGGACATCATCATGGAGCAC  
 TCTGTCACTGATGGGCTCCTGGGGCCCCGGTGCCAAGCCCAGCGTGGAGGCGGCAGGCACGGG  
 ATGCTGCTTGTGCTACCCCAAT**TAG**GGGCCAGCCTGCCATGGCAGCTCCAACAAGGACCGTC  
 TGGGATAGCACCTCTCAGTCAACATCGTGGGGCTGGACAGGGCTGCGGCCCTCTGCCCACA  
 CTCAGTACTGACCAAAGCCAGGGCTGTGTGTGCCTGTGTGTAGGTCCCACGGCCTCTGCCTC  
 CCCAGGGAGCAGAGCCTGGGCCTCCCCTAAGAGGCTTTCCCCGAGGCAGCTCTGGAATCTGT  
 GCCCACCTGGGGCCTGGGGAACAAGGCCCTCCTTTCTCCAGGCCTGGGCTACAGGGGAGGGA  
 GAGCCTGAGGCTCTGCTCAGGGCCCATTTTCATCTCTGGCAGTGCCTTGGCGGTGGTATTCAA  
 GGCAGTTTTGTAGCACCTGTAATTGGGGAGAGGGAGTGTGCCCCCTCGGGGCAGGAGGGAAGG  
 GCATCTGGGGAAGGGCAGGAGGGAAGAGCTGTCCATGCAGCCACGCCCATGGCCAGGTGGC  
 CTCTTCTCAGCCTCCCAGGTGCCTTGAGCCCTCTTGCAAGGGCGGCTGCTTCCTTGAGCCTA  
 GTTTTTTTTTACGTGATTTTTGTAACATTCATTTTTTTGTACAGATAACAGGAGTTTCTGAC  
 TAATCAAAGCTGGTATTTCCCCGCATGTCTTATTCTTGCCCTTCCCCCAACCAGTTTGTAA  
 TCAAACAATAAAAAACATGTTTTGTTTTGTTTTTAAAAA

**FIGURE 70**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64863

><subunit 1 of 1, 294 aa, 1 stop

><MW: 33211, pI: 5.35, NX(S/T): 3

MPRGDSEQVRYCARFSYLWLKFSLLIYSTVFWLIGALVLSVGIYAEVERQKYKTLES AFLAP  
AIILILLGVVMFMVSFIGVLASLRDNLVLLQAFMYILGICLIMELIGGVVALTFRNQTIDFL  
NDNIRRGIE NYDDLDLDFKNIMDFVQKKFKCCGGEDYRDWSKNQYHDCSAPGPLACGV PYTCC  
IRNTTEVVNTMCGYKTIDKERFSVQDVIYVRGCTNAV IIFMDNYTIMACILLGILLPQFLG  
VLLTLLYITRVEDIIMEHSVTDG LLGPGAKPSVEAAGTGCCLCYPN

**Signal peptide:**

amino acids 1-44

**Transmembrane domains:**

amino acids 22-42, 57-85, 93-116, 230-257

## FIGURE 71

GAGGAGCGGGCCGAGGACTCCAGCGTGCAGGTCTGGCATCCTGCACTTGGCTGCCCTCTGA  
CACCTGGGAAGATGGCCGGCCCGTGGACCTTCACCCTTCTCTGTGGTTTGGCTGGCAGCCACC  
TTGATCCAAGCCACCCTCAGTCCCAGTGCAGTTCTCATCCTCGGCCCAAAGTCATCAAAGA  
AAAGCTGACACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGC  
TCAGTGCCATGCGGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACC  
GTCCTGAAGCACATCATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAA  
GCCCTCGGCCAATGACCAGGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCA  
ACACGCCCCCTGGTCAAGACCATCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATC  
CGCATGGACACCAGTGCAAGTGGCCCCACCCGCCTGGTCCTCAGTGACTGTGCCACCAGCCA  
TGGGAGCCTGCGCATCCAAGTCTGTATAAGCTCTCCTTCTGGTGAACGCCTTAGCTAAGC  
AGGTCATGAACCTCCTAGTGCCATCCCTGCCCAATCTAGTGAAAAACCAGCTGTGTCCCGTG  
ATCGAGGCTTCCCTTCAATGGCATGTATGCAGACCTCCTGCAGCTGGTGAAGGTGCCCATTTT  
CCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTATCCTGCCATCAAGGGTGACACCATT  
AGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAAAGGTGACCAAGTGGTTCAATAAC  
TCTGCAGCTTCCCTGACAATGCCACCCCTGGACAACATCCCGTTCAGCCTCATCGTGAGTCA  
GGACGTGGTGAAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGAATTCATGGTCCTGTTGG  
ACTCTGTGCTTCCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGCTGATCAATGAAAAG  
GCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGACACTCCCGAGTT  
TTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTTTCCCTCCA  
GTGAAGCCCTCCGCCCTTTGTTACCCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGTTTAC  
ACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTGAT  
GAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCC  
ACTCCATCCTGCTGCCGAACCAGAAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTG  
AAGGCCTTGGGATTGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTTGTGCTTACTCC  
AGCCTCCTTGTGGAAACCCAGCTCTCCTGTCTCCCAGTGAAGACTTGGATGGCAGCCATCAG  
GGAAGGCTGGGTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAAT  
CAATAAACACTTGCCTGTGAAAAA

**FIGURE 72**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64881

><subunit 1 of 1, 484 aa, 1 stop

><MW: 52468, pI: 7.14, NX(S/T): 3

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIEKLTQELKDHNATSILQQLPLLSAM  
REKPAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPL  
VKTIVEFHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMN  
LLVPSLPNLVKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYL  
GAKLLDSQGKVTWKFNNSAASLTMPITLDNIPFSLIVSQDVVKAAVAVALSPPEEFMVLLDSVL  
PESAHRLKSSIGLINEKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEAL  
RPLFTLGIEASSEAQFYTKGDQLILNLNLISSDRIQLMNSGIGWFQPDVLKNIITEIIHSIL  
LPNQNGKLRSGVPVSLVKALGFEEAESSLTkdALVLTpasLWKpSSpVSQ

**Important features of the protein:****Signal peptide:**

amino acids 1-21

**N-glycosylation sites.**

amino acids 48-51, 264-267, 401-404

**Glycosaminoglycan attachment site.**

amino acids 412-415

**LBP / BPI / CETP family proteins.**

amino acids 407-457

### FIGURE 73

GAGCGAAC**ATG**GCAGCGCGTTGGCGGTTTTGGTGTGTCTCTGTGACCATGGTGGTGGCGCTGCTCATCGTTTTGCGACGTTCCCTCAGCCTCTGCCCAAAGAAAGAAGGAGATGGTGTATCTGAAAGGTTAGTCAGCTGATGGAATGGACTAACAAAAGACCTGTAATAAGAATGAATGGAGACAAGTCCGTCGCCTTGTGAAAGCCCCACCGAGAAATTACTCCGTTATCGTCATGTTCACTGCTCTCCAACCTGCATAGACAGTGTGTGTTTTGCAAGCAAGCTGATGAAGAATTCAGATCCTGGCAAACCTCCTGGCGATACTCCAGTGCATTACCAACAGGATATTTTTTGGCATGGTGGATTTTTGATGAAGGCTCTGATGTATTTTCTGATGCTAAACATGAATTCAGCTCCAACCTTTCATCAACTTTTCTGCAAAAGGGAAACCCAAACGGGGTGATACATATGAGTTACAGGTGCGGGGTTTTTTCAGCTGAGCAGATTGCCCGGTGGATCGCCGACAGAACTGATGTCAATATTAGAGTGATTAGACCCCAAATTATGCTGGTCCCTTATGTTGGGATTGCTTTTGGCTGTTATTGGTGGACTTGTGTATCTTCGAAGAAGTAATATGGAATTTCTCTTTAATAAAACTGGATGGGCTTTTTGCAGCTTTGTGTTTTGTGCTTGCTATGACATCTGGTCAAATGTGGAACCATATAAGAGGACCACCATATGCCATAAGAATCCCCACACGGGACATGTGAATTATATCCATGGAAGCAGTCAAGCCCAGTTTGTAGCTGAAACACACATTGTTCTTCTGTTAATGGTGGAGTTACCTTAGGAATGGTGGCTTTTATGTGAAGCTGCTACCTCTGACATGGATATTGGAAAGCGAAAGATAATGTGTGTGGCTGGTATTGACTTGTTGTATTATTCTTCAGTTGGATGCTCTCTATTTTTTAGATCTAAATATCATGGCTACCCATACAGCTTTCTGATGAGT**TAA**AAAGGTCCCAGAGATATATAGACACTGGAGTACTGGAAATTGAAAAACGAAAATCGTGTGTGTTTTGAAAAGAAGAATGCAACTTGTATATTTTGTATTACTCTTTTTTTTCAAGTGATTTAAATAGTTAATCATTTAACCAGAAAGATGTGTAGTGCCTTACAAGCAATCCTCTGTCAAAATCTGAGGTATTTGAAAATAATTATCCTCTTAACCTTCTCTTCCCAGTGAACTTTATGGAACATTTAATTTAGTACAATTAAGTATATTATAAAAAATTGTAAAACTACTACTTTGTTTTAGTTAGAACAAAGCTCAAAACTACTTTAGTTAACTTGGTCATCTGATTTTATATTGCCTTATCCAAAGATGGGGAAAGTAAGTCTGACCAGGTGTTCCACATATGCCGTTACAGATAACTACATTAGGAATTCATTCTTAGCTTCTTCATCTTTGTGTGGATGTGTATACTTTACGCATCTTTCCTTTTGAGTAGAGAAATTATGTGTGTCATGTGGTCTTCTGAAAATGGAACACCATTCTTCAGAGCACACGTCTAGCCCTCAGCAAGACAGTTGTTTCTCCTCCTCCTTGCATATTTCCCTACTGCGCTCCAGCCTGAGTGATAGAGTGAGACTCTGTCTCAAAAAAAGGATCTCTAAATACAGGATTATAATTTCTGCTTGAGTATGGTGTAACTACCTTGTATTTAGAAAGATTTTCTGATTTTCAATTCATCTCCTTAGTTTTCTTTTAAAGGTGACCCATCTGTGATAAAAAATAGCTTAGTGCTAAAATCAGTGTAACCTTATACATGGCCTAAAATGTTTCTACAAATTAGAGTTGTCACTTATTCCATTTGTACCTAAGAGAAAAATAGGCTCAGTTAGAAAAGGACTCCCTGGCCAGGCGCAGTGACTTACGCCTGTAATCTCAGCACTTTGGGAGGCCAAGGCAGGCAGATCACGAGGTGAGGAGTTCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATATAAAAAATTAGCTGGGTGTGGTGGCAGGAGCCTGTAATCCCAGCTACACAGGAGGCTGAGGCACGAGAATCACTTGAACCTCAGGAGATGGAGGTTTTCAGTGAGCCGAGATCACGCCACTGCACTCCAGCCTGGCAACAGAGCGAGACTCCATCTCAAAAAAAAAAAAAA

### FIGURE 74

MAARWRFWCVSVTMVVALLIIVCDVPSASAQRKKEMVLSEKVSQ LMEWTNKR P VIRMNGDKFR  
RLVKAPPRNYSVIVMF TALQLHRQCVVCKQADEEFQILANSWRYSSAFTNRIF FAMVDFDEG  
SDVFQMLNMNSAPTFINFP AKGKPKRGDTYELQVRGFSAEQIARWIADRTDVNIRVIRPPNY  
AGPLMLGLLLAVIGGLVYLRRSNMEFLFNKTGWAF AALCFVLAMTSGQMWNHIRGPPYAHKN  
PHTGHVNYIHGSSQAQFVAETHIVLLFNGGVTLGMVLLCEAATSMDDIGKRKIMCVAGIGLV  
VLFFSWMLSIFRSKYHGYPYSFLMS

Signal peptide:

amino acids 1-29

**Transmembrane domains:**

amino acids 183-205, 217-237, 217-287, 301-321

**FIGURE 75**

AAGCAACCAAACCTGCAAGCTTTGGGAGTTGTTTCGCTGTCCCTGCCCTGCTCTGCTAGGGAGA  
 GAACGCCAGAGGGAGGCGGCTGGCCCGGCGGCAGGCTCTCAGAACCGCTACCGGCG**ATG**CTA  
 CTGCTGTGGGTGTCTGGTGGTCGCAGCCTTGGCGCTGGCGGTACTGGCCCCCGGAGCAGGGGA  
 GCAGAGGCGGAGAGCAGCCAAAGCGCCCAATGTGGTGCTGGTCGTGAGCGACTCCTTCGATG  
 GAAGGTAAACATTTTCATCCAGGAAGTCAGGTAGTGAAACTTCCTTTTATCAACTTTATGAAG  
 ACACGTGGGACTTCCTTTCTGAATGCCTACACAAACTCTCCAATTTGTTGCCCATCACGCGC  
 AGCAATGTGGAGTGGCCTCTTCACTCACTTAACAGAATCTTGGAATAATTTTAAGGGTCTAG  
 ATCCAAATTATACAACATGGATGGATGTCATGGAGAGGCATGGCTACCGAACACAGAAATTT  
 GGGAACTGGACTATACTTCAGGACATCACTCCATTAGTAATCGTGTGGAAGCGTGGACAAG  
 AGATGTTGCTTTCTTACTCAGACAAGAAGGCAGGCCCATGGTTAATCTTATCCGTAACAGGA  
 CTAAAGTCAGAGTGATGGAAAGGGATTGGCAGAATACAGACAAAGCAGTAAACTGGTTAAGA  
 AAGGAAGCAATTAATTACACTGAACCATTTGTTATTTACTTGGGATTAAATTTACCACACCC  
 TTACCCTTCACCATCTTCTGGAGAAAAATTTGGATCTTCAACATTTTACACATCTCTTTATT  
 GGCTTGAAAAAGTGTCTCATGATGCCATCAAAATCCCAAAGTGGTCACCTTTGTCAGAAATG  
 CACCCTGTAGATTATTACTCTTCTTATACAAAAAACTGCACTGGAAGATTTACAAAAAAGA  
 AATTAAGAATATTAGAGCATTTTATTATGCTATGTGTGCTGAGACAGATGCCATGCTTGGTG  
 AAATTATTTTGGCCCTTCATCAATTAGATCTTCTTCAGAAAACTATTGTCATATACTCCTCA  
 GACCATGGAGAGCTGGCCATGGAACATCGACAGTTTATAAAATGAGCATGTACGAGGCTAG  
 TGCACATGTTCCGCTTTTGATGATGGGACCAGGAATTAAAGCCGGCCTACAAGTATCAAATG  
 TGGTTTCTCTTGTGGATATTTACCCTACCATGCTTGATATTGCTGGAATTCCTCTGCCTCAG  
 AACCTGAGTGGATACTCTTTGTTGCCGTTATCATCAGAAACATTTAAGAATGAACATAAAGT  
 CAAAAACCTGCATCCACCCTGGATTCTGAGTGAATTCATGGATGTAATGTGAATGCCTCCA  
 CCTACATGCTTCGAACTAACCCTGGAAATATATAGCCTATTTCGGATGGTGCATCAATATTG  
 CCTCAACTCTTTGATCTTTCTCGGATCCAGATGAATTAACAAATGTTGCTGTAAAAATTTCC  
 AGAAATTACTTATTCTTTGGATCAGAAGCTTCATTCCATTATAAACTACCCTAAAGTTTCTG  
 CTTCTGTCCACCAGTATAATAAAGAGCAGTTTATCAAGTGGAAACAAAGTATAGGACAGAAT  
 TATTCAAACGTTATAGCAAATCTTAGGTGGCACCAAGACTGGCAGAAGGAACCAAGGAAGTA  
 TGAAAATGCAATTGATCAGTGGCTTAAACCCCATATGAATCCAAGAGCAGTT**TGA**ACAAAAA  
 GTTTAAAAATAGTGTTCTAGAGATACATATAAATATATTACAAGATCATAATTATGTATTTT  
 AAATGAAACAGTTTTAATAATTACCAAGTTTGGCCGGGCACAGTGGCTCACACCTGTAATC  
 CCAGGACTTTGGGAGGCTGAGGAAAGCAGATCACAAAGGTCAAGAGATTGAGACCATCCTGGC  
 CAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCTGGGCGCGGTGGTGCACA  
 CCTATAGTCTCAGCTACTCAGAGGCTGAGGCAGGAGGATCGCTTGAACCCGGGAGGCAGCAG  
 TTGCAGTGAGCTGAGATTGCGCCACTGTACTCCAGCCTGGCAACAGAGTGAGACTGTGTGCGC  
 AAAAAATAAAAAATAAATAAATAAATTACCAATTTTTCATTATTTTGTAAAGATGTAGTG  
 TATTTTAAGATAAAATGCCAATGATTATAAAATCACATATTTTCAAAATGGTTATTATTTA  
 GGCCTTTGTACAATTTCTAACAATTTAGTGGAAGTATCAAAGGATTGAAGCAAATACTGTA  
 ACAGTTATGTTCTTTAAATAATAGAGAATATAAAATATTGTAATAATATGTATCATAAAT  
 AGTTGTATGTGAGCATTTGATGGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
 AA

**FIGURE 76**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64885

<subunit 1 of 1, 536 aa, 1 stop

<MW: 61450, pI: 9.17, NX(S/T): 7

MLLLWVSVVAALALAVLAPGAGEQRRRAAKAPNVVLVVSDFSFDGRLTFHPGSQVVKLPFINE  
MKTRGTSFLNAYTNSPICCPSSRAAMWSGLFTHLTESWNNFKGLDPNYTTWMDVMERHGYRTQ  
KFGKLDYTSGHHSISNRVEAWTRDVAFLLRQEGRPMVNLIRNRTKVRVMERDWQNTDKAVNW  
LRKEAINYTEPFVIYLG LNLPHYPSPSSGENFGSSTFHTSLYWLEKVSHDAIKIPKWSPLS  
EMHPVDYYSYTKNCTGRFTKKEIKNIRAFYYAMCAETDAMLGEIILALHQDLLQKTIVIIY  
SSDHGELAMEHRQFYKMSMYEASAHVPLLMMGPGIKAGLQVSNVVS LVDIYPTMLDIAGIPL  
PQNLSGYSLPLSSETFKNEHKVKNLHPPWILSEFHGCNVNASTYMLRTNHWKYIAYS DGAS  
ILPQLFDLSSDPDEL TNVAVKFPEITYSLDQKLHSIINYPKVSASVHQYNKEQFIKWKQSIG  
QNYSNVIANLRWHQDWQKEPRKYENAIQWLKTHMNPRAV

**Important features:****Signal peptide:**

amino acids 1-15

**N-glycosylation sites.**

amino acids 108-111, 166-169, 193-196, 262-265, 375-378, 413-416,  
498-501

**Sulfatases proteins:**

amino acids 286-315, 359-369, 78-97



**FIGURE 77**

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAG  
 AGCTTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCC**ATG**  
 GCCTCTCTTGGCCTCCAACCTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACT  
 GGTGCCATGCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAG  
 CAGTTGGCTTCTCCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAG  
 TGTGACATCTATAGCACCTTCTGGGCCTGCCCGCTGACATCCAGGCTGCCCAGGCCATGAT  
 GGTGACATCCAGTGCAATCTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCA  
 CAGTCTTCTGCCAGGAATCCCGAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTC  
 ATCCTTGGAGGCCTCCTGGGATTCATTCCCTGTTGCCTGGAATCTTCATGGGATCCTACGGGA  
 CTTCTACTCACCACTGGTGCCTGACAGCATGAAATTTGAGATTGGAGAGGCTCTTTACTTGG  
 GCATTATTTCTTCCCTGTTCTCCCTGATAGCTGGAATCATCCTCTGCTTTTCCTGCTCATCC  
 CAGAGAAATCGCTCCAACCTACTACGATGCCTACCAAGCCCAACCTCTTGCCACAAGGAGCTC  
 TCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGAGTTCAATTCCTACAGCCTGACAGGGT  
 ATGTG**TGA**AGAACCAGGGGCCAGAGCTGGGGGGTGGCTGGGTCTGTGAAAAACAGTGGACAG  
 CACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGTGTGAGAAGGTGCTGCTGAGG  
 ATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGGGCTAGTGTAACAGCATG  
 CAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCTCACCTTGCTGCTC  
 CCCTGCCCTAAGTCCCCAACCCTCAACTTGAAACCCCATTCCTTAAGCCAGGACTCAGAGG  
 ATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACATCCCACTG  
 ACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATTGCTGG  
 GGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTCCC  
 TCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCACTCTTGTTATGACTCCACAGTGTCC  
 AGACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAG  
 GATGCAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

**FIGURE 78**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64886
><subunit 1 of 1, 230 aa, 1 stop
><MW: 24549, pI: 8.56, NX(S/T): 1
MASLGLQLVGYILGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGIT
QCDIYSTLLGLPADIQAAQAMMVTSSAIISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVF
FILGGLLGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCS
SQNRNSNYDAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV
```

**Important features of the protein:****Signal peptide:**

amino acids 1-24

**Transmembrane domains:**

amino acids 82-102, 117-140, 163-182

**N-glycosylation site.**

amino acids 190-193

**PMP-22 / EMP / MP20 family proteins.**

amino acids 46-59

**FIGURE 79**

GCACTGCTGCTGTCCCATCAGCTGCTCTGAAGCTCCATGGTGCCCAGAATCTTCGCTCCTGC  
TTATGTGTCAGTCTGTCTCCTCCTCTTGTGTCCAAGGGAAGTCATCGCTCCCGCTGGCTCAG  
AACCATGGCTGTGCCAGCCGGCACCCAGGTGTGGAGACAAGATCTACAACCCCTTGGAGCAG  
TGCTGTTACAATGACGCCATCGTGTCCCTGAGCGAGACCCGCCAATGTGGTCCCCCCTGCAC  
CTTCTGGCCCTGCTTTGAGCTCTGCTGTCTTGATTCTTTGGCCTCACAAACGATTTTGTTG  
TGAAGCTGAAGGTTTCAGGGTGTGAATTCCCAGTGCCACTCATCTCCCATCTCCAGTAAATGT  
GAAAGCAGAAGACGTTTTCCCTTGAGAAAGACATAGAAAGAAAATCAACTTTCACTAAGGCATC  
TCAGAAACATAGGCTAAGGTAATATGTGTACCAGTAGAGAAGCCTGAGGAATTTACAAAATG  
ATGCAGCTCCAAGCCATTGTATGGCCCATGTGGGAGACTGATGGGACATGGAGAATGACAGT  
AGATTATCAGGAAATAAATAAAGTGGTTTTTCCAATGTACACACCTGTAAAA

## **FIGURE 80**

MVPRI FAPAYVSVCLLLLCPREVIAPAGSEPWLCQPAPRCGDKIYNPLEQCCYNDAIVSLSE  
TRQCGPPCTFWPCFELCCLDSFGLTND FVVKLVQGVNSQCHSSPISSKCESRRRFP

**Signal peptide:**

amino acids 1-25

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

CTCCACTGCAACCCAGAGCCATGGCTCCCCGAGGCTGCATCGTAGCTGTCTTTGCCATT  
TTCTGCATCTCCAGGCTCCTCTGCTCACACGGAGCCCCAGTGGCCCCCATGACTCCTTACCT  
GATGCTGTGCCAGCCACACAAGAGATGTGGGGACAAGTTCTACGACCCCCTGCAGCACTGTT  
GCTATGATGATGCCGTCTGTGCCCTTGGCCAGGACCCAGACGTGTGGAAACTGCACCTTCAGA  
GTCTGCTTTGAGCAGTGCTGCCCCCTGGACCTTCATGGTGAAGCTGATAAACCAGAACTGCGA  
CTCAGCCCGGACCTCGGATGACAGGCTTTGTCTGCAGTGTCTAGCTAATGGAACATCAGGGGAA  
CGATGACTCCTGGATTCTCCTTCCTGGGTGGGCCTGGAGAAAGAGGCTGGTGTACCTGAGA  
TCTGGGATGCTGAGTGGCTGTTTGGGGGCCAGAGAAACACACACTCAACTGCCCACTTCATT  
CTGTGACCTGTCTGAGGCCACCCTGCAGCTGCCCTGAGGAGGCCACAGGTCCCCTTCTAG  
AATTCTGGACAGCATGAGATGCGTGTGCTGATGGGGGCCAGGGACTCTGAACCCTCCTGAT  
GACCCCTATGGCCAACATCAACCCGGCACCACCCCAAGGCTGGCTGGGGAACCCTTCACCCCT  
TCTGTGAGATTTTCCATCATCTCAAGTTCTCTTCTATCCAGGAGCAAAGCACAGGATCATAA  
TAAATTTATGTACTTTATAAATGAAAA

## **FIGURE 82**

MAPRGCIVAVFAIFCISRLLC SHGAPVAPMTPYLMLCQPHKRCGDKFYDPLQHCCYDDAVVP  
LARTQTCGNCTFRVCFEQCCPWTFMVKLINQNCDSARTSDDRRCRSVS

**Signal peptide:**

amino acids 1-24

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250

**FIGURE 83**

GGGGGCGGGTGCCTGGAGCACGGCGCTGGGGCCGCCCAGCGCTCACTCGCTCGCACTCAG  
 TCGCGGGAGGCTTCCCCGCGCCGGCCGCGTCCCGCCCGCTCCCCGGCACCAGAAAGTTCCCTCT  
 GCGCGTCCGACGGCGACATGGGCGTCCCCACGGCCCTGGAGGCCGGCAGCTGGCGCTGGGGA  
 TCCCTGCTCTTCGCTCTCTTCCTGGCTGCGTCCCTAGGTCCGGTGGCAGCCTTCAAGGTCGC  
 CACGCCGTATTCCCTGTATGTCTGTCCCGAGGGGCAGAACGTCACCCTCACCTGCAGGCTCT  
 TGGGCCCTGTGGACAAAGGGCACGATGTGACCTTCTACAAGACGTGGTACCGCAGCTCGAGG  
 GCGAGGTGCAGACCTGCTCAGAGCGCCGGGCCATCCGCAACCTCACGTTCCAGGACCTTCA  
 CCTGCACCATGGAGGCCACCAGGCTGCCAACACCAGCCACGACCTGGCTCAGCGCCACGGGC  
 TGGAGTCGGCCTCCGACCACCATGGCAACTTCTCCATCACCATGCGCAACCTGACCCTGCTG  
 GATAGCGGCCTCTACTGCTGCCTGGTGGTGGAGATCAGGCACCACCCTCGGAGCACAGGGT  
 CCATGGTGGCATGGAGCTGCAGGTGCAGACAGGCAAAGATGCACCATCCAACTGTGTGGTGT  
 ACCCATCCTCCTCCCAGGATAGTGAAAACATCACGGCTGCAGCCCTGGCTACGGGTGCCTGC  
 ATCGTAGGAATCCTCTGCCTCCCCCTCATCCTGCTCCTGGTCTACAAGCAAAGGCAGGCAGC  
 CTCCAACCGCCGTGCCAGGAGCTGGTGC GGATGGACAGCAACATTCAAGGGATTGAAAACC  
 CCGGCTTTGAAGCCTCACCACCTGCCAGGGGATACCCGAGGCCAAAGTCAGGCACCCCCTG  
 TCCTATGTGGCCCAGCGGCAGCCTTCTGAGTCTGGGCGGCATCTGCTTTCGGAGCCCAGCAC  
 CCCCCTGTCTCCTCCAGGCCCGGAGACGTCTTCTTCCCATCCCTGGACCCTGTCCCTGACT  
 CTCCAAACTTTGAGGTCATCTAGCCCAGCTGGGGGACAGTGGGCTGTTGTGGCTGGGTCTGG  
 GGCAGGTGCATTTGAGCCAGGGCTGGCTCTGTGAGTGGCCTCCTTGGCCTCGGCCCTGGTTC  
 CCTCCCTCCTGCTCTGGGCTCAGATACTGTGACATCCCAGAAGCCCAGCCCCTCAACCCCTC  
 TGGATGCTACATGGGGATGCTGGACGGCTCAGCCCCTGTTCCAAGGATTTTGGGGTGCTGAG  
 ATTCTCCCCTAGAGACCTGAAATTACCAGCTACAGATGCCAAATGACTTACATCTTAAGAA  
 GTCTCAGAACGTCCAGCCCTTCAGCAGCTCTCGTTCTGAGACATGAGCCTTGGGATGTGGCA  
 GCATCAGTGGGACAAGATGGACACTGGGCCACCCTCCCAGGCACCAGACACAGGGCACGGTG  
 GAGAGACTTCTCCCCGTGGCCGCCTTGGCTCCCCCGTTTTGCCCGAGGCTGCTCTTCTGTC  
 AGACTTCCTCTTTGTACCACAGTGGCTCTGGGGCCAGGCCTGCCTGCCACTGGCCATCGCC  
 ACCTTCCCCAGCTGCCTCCTACCAGCAGTTTCTCTGAAGATCTGTCAACAGGTAAAGTCAAT  
 CTGGGGCTTCCACTGCCTGCATTCCAGTCCCAGAGCTTGGTGGTCCCGAAACGGGAAGTAC  
 ATATTGGGGCATGGTGGCCTCCGTGAGCAAATGGTGTCTTGGGCAATCTGAGGCCAGGACAG  
 ATGTTGCCCCACCCACTGGAGATGGTGTGAGGGAGGTGGGTGGGGCCTTCTGGGAAGGTGA  
 GTGGAGAGGGGCACCTGCCCCCGCCCTCCCATCCCCTACTCCCCTGCTCAGCGCGGGCC  
 ATTGCAAGGGTGCCACACAATGTCTTGTCCACCCTGGGACACTTCTGAGTATGAAGCGGGAT  
 GCTATTAAAACTACATGGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGA

**FIGURE 84**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64897

><subunit 1 of 1, 311 aa, 1 stop

><MW: 33908, pI: 6.87, NX(S/T): 6

MGVPTALEAGSWRWGSLLFALFLAASLGPVAAFKVATPYSLYVCPEGQNVTLTCRLLGPVDK  
GHDVTFYKTWYRSSRGEVQTCSERRPIRNLTFQDLHLHHGGHQAANTSHDLAQRHGLESASD  
HHGNFSITMRNLTLLDGLYCCLVVEIRHHHSEHRVHGAMELQVQTGKDAPSNCVVYPSSSQ  
DSENITAAALATGACIVGILCLPLILLLVYKQRQAASNRRAQELVRMDSNIQGIENPGFEAS  
PPAQGIPEAKVRHPLSYVAQRQPSESGRHLLSEPSTPLSPPGPGDVFFPSLDPVPDSPNFEVI

**Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 190-216

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



**FIGURE 85**

CCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCT  
TTCCCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCCACCTTAGACCTCCCTTCCTGCCCCTCC  
TTTCTGCCCCACCGCTGCTTCCTGGCCCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGG  
GGTCTGTGGGTTGATCTGTGGCCCCCTGTGCCTCCGTGTCTTTTTCGTCTCCCTTCCTCCCGA  
CTCCGCTCCCGGACCAGCGGCCTGACCCTGGGGAAAGGATGTTCCCGAGGTGAGGGTCCTC  
TCCTCCTTGCTGGGACTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCC  
AGACATGTTCTGCCTTTTCCATGGGAAGAGATACTCCCCCGCGAGAGCTGGCACCCCTACT  
TGGAGCCACAAGGCCTGATGTACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGT  
TGTTACCGCCTCCACTGTCCGCTGTCCACTGCCCCCAGCCTGTGACGGAGCCACAGCAATG  
CTGTCCCAAGTGTGTGGAACCTCACACTCCCTCTGGACTCCGGGGCCCCACCAAAGTCCTGCC  
AGCACAACGGGACCATGTACCAACACGGAGAGATCTTCAGTGCCCATGAGCTGTTCCCTCC  
CGCCTGCCCAACCAGTGTGTCTCTGCAGCTGCACAGAGGGCCAGATCTACTGCGGCCTCAC  
AACCTGCCCCGAACCAGGCTGCCCAGCACCCCTCCCACTGCCAGACTCCTGCTGCCAAGCCT  
GCAAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACAGTGTGCAGTCGCTCCATGGGGTG  
AGACATCCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAGAGAGGCCCGGGCACCCCAGC  
CCCCACTGGCCTCAGCGCCCCTCTGAGCTTCATCCCTCGCCACTTCAGACCCAAGGGAGCAG  
GCAGCACAACCTGTCAAGATCGTCCTGAAGGAGAAACATAAGAAAGCCTGTGTGCATGGCGGG  
AAGACGTACTCCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGCCCCTTGCCCTG  
CATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCCCACCGAGT  
ACCCCTGCCGTCACCCCGAGAAAGTGGCTGGGAAGTGCTGCAAGATTTGCCCAGAGGACAAA  
GCAGACCCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCCGGGTCT  
CGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGG  
CCTCGGACTTGGTGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAG  
AGAGGTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGA  
AAGTCAGGAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCCAC  
GAAGGTCACTGGAACGTCTTCCTAGCCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGA  
CAAAGTGACCAAGACATAACAAAGACCTAAACAGTTGCAGATATGAGCTGTATAATTGTTGTT  
ATTATATATTAATAAATAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 86**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64902

><subunit 1 of 1, 451 aa, 1 stop

><MW: 49675, pI: 7.15, NX(S/T): 1

MVPEVRVLSSLLGLALLWFPLDSHARARPD MFCLFHGKRYSPGESWHPYLEPQGLMYCLRCT  
CSEGAHVSCYRLHCPPVHCPQPVTEPQQCCPKVEPHTPSGLRAPPKSCQHNGTMYQHGEIF  
SAHELFPSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEED  
SVQSLHGVRHPQDPCSSDAGRKRGPPTAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKH  
KKACVHGGKTYSHGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPT EYPCRHPEKVAGKC  
CKICPEDKADPGHSEISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKL  
KDEETEAQRGEVPGPRPHSQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSPDPG  
AEGHGQSRQSDQDITKT

**Signal peptide:**

amino acids 1-25

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 87**

CTAGCCTGCGCCAAGGGGTAGTGAGACCGCGCGGCAACAGCTTGCGGGCTGCGGGGAGCTCCC  
GTGGGCGCTCCGCTGGCTGTGCAGGCGGCC**ATG**GATTTCCTTGCGGAAAATGCTGATCTCAGT  
CGCAATGCTGGGCGCAGGGGCTGGCGTGGGCTACGCGCTCCTCGTTATCGTGACCCCGGGAG  
AGCGGCGGAAGCAGGAAATGCTAAAGGAGATGCCACTGCAGGACCCAAGGAGCAGGGAGGAG  
GCGGCCAGGACCCAGCAGCTATTGCTGGCCACTCTGCAGGAGGCAGCGACCACGCAGGAGAA  
CGTGGCCTGGAGGAAGAAGTGGATGGTTGGCGGCGAAGGCGGCGCCAGCGGGAGGTCACCG**T**  
**GAG**ACCGGACTTGCCCTCCGTGGGCGCCGGACCTTGGCTTGGGCGCAGGAATCCGAGGCAGCC  
TTTCTCCTTCGTGGGCCCAGCGGAGAGTCCGGACCGAGATACCATGCCAGGACTCTCCGGGG  
TCCTGTGAGCTGCCGTGCGGTGAGCACGTTTCCCCCAAACCCTGGACTGACTGCTTTAAGGT  
CCGCAAGGCGGGCCAGGGCCGAGACGCGAGTCGGATGTGGTGAAGTAAAGAACCAATAAAA  
TCATGTTCTCCAA  
AAAAAAAAAAAA

## **FIGURE 88**

MDSLRLKMLISVAMLGAGAGVGYALLVIVTPGERRKQEMLKEMPLQDPRSREEAARTQQLLLA  
TLQEAATTQENVAWRKNWMVGGEGGASGRSP

**Signal peptide:**

amino acids 1-18

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

## FIGURE 89

[illegible]

**FIGURE 90**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64905

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46038, pI: 6.50, NX(S/T): 2

MGPSTPLLILFLLSWGPLQGQHHLVEYMERRLAALAEERLAQCQDQSSRHAAELRDFKNKM  
LPLLEVAEKEREALRTEADTISGRVDRLEREVDYLETQNPALPCVEFDEKVTGGPGTKGKGR  
RNEKYDMVTDCGYTISQVRSMKILKRFGGPAGLWTKDPLGQTEKIYVLDGTQNDTAFVFPRL  
RDFTLAMAARKASRVVPFPWVG TGQLVYGGFLYFARRPPGRPGGGGEMENTLQLIKFHLAN  
RTVVDSSVFPAEGLIPPYGLTADTYIDLVADEEGLWAVYATREDDRHLCLAKLDPQTLDEQ  
QWDTPCPRENAEAAAFVICGTLVYVYNTRPASRARIQCSFDASGTLTPERAALPYFPRRYGAH  
ASLRYNPRERQLYAWDDGYQIVYKLEM RKKEEEV

**Important features:****Signal peptide:**

amino acids 1-21

**N-glycosylation sites.**

amino acids 177-180, 248-251

**FIGURE 91**

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCACGCAG  
 AGCCTCTCCGTGGCTTCCGCACCTTGAGCATTAGGCCAGTTTCCTCTTCTCTCTAATCCAT  
 CCGTCACCTCTCCTGTCATCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCC**ATGG**  
 CTCTCATGCTCAGTTTGGTTCTGAGTCTCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTT  
 GGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGGAGGACGCAGCATTCTCCTGTTTCCTGTC  
 TCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGGCCAGTTCTCTAGCGTGG  
 TCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATCAAGGCAGG  
 ACAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTAC  
 TGTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCCAGTCTTACTACCAGAAGGCCA  
 TCTGGGAGCTACAGGTGTCAGCACTGGGCTCAGTTCCCTCTCATTTCATCACGGGATATGTT  
 GATAGAGACATCCAGCTACTCTGTGAGTCCTCGGGCTGGTTCCCCCGGCCACAGCGAAGTG  
 GAAAGGTCCACAAGGACAGGATTTGTCCACAGACTCCAGGACAAACAGAGACATGCATGGCC  
 TGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAACGCCGGGAGCATATCCTGTTCCATG  
 CGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAGGAGATACCTTTTTTCGA  
 GCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTATTTTTTG  
 GCATTGTTGGACTGAAGATTTTCTTCTCCAAATTCCAGTGGAAAATCCAGGCGGAACTGGAC  
 TGGAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACACGCAGTGGAGGTGAC  
 TCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTGAACCCATA  
 GAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCT  
 TCTCAGAGTTTCCAAGCAGGGAAACATTACTGGGAGGTGGACGGAGGACACAATAAAAGGTG  
 GCGCGTGGGAGTGTGCCGGGATGATGTGGACAGGAGGAAGGAGTACGTGACTTTGTCTCCCG  
 ATCATGGGTACTGGGTCCTCAGACTGAATGGAGAACATTTGTATTTACATTAAATCCCCGT  
 TTTATCAGCGTCTTCCCCAGGACCCACCTACAAAAATAGGGGTCTTCCTGGACTATGAGTG  
 TGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTCGGT  
 TTGAAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAAATGGAACCTCCC  
 ATAGTCATCTGCCCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGC  
 AATCCCAGAGACAAGCAACAGTGAGTCCTCCTCACAGGCAACCACGCCCTTCCTCCCCAGGG  
 GTGAAATG**TAG**GATGAATCACATCCACATTCTTCTTTAGGGATATTAAGGTCTCTCTCCA  
 GATCCAAAGTCCCGCAGCAGCCGGCCAAGGTGGCTTCCAGATGAAGGGGGACTGGCCTGTCC  
 ACATGGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGGAGGGAAGAAGGCTGACATTACATTT  
 AGTTTGCTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAGAACCG  
 TCAGGAATTCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGC  
 TTAGATCTTATTGATGACAGAGTGTATCCTAATGGTTTGTTCATTATATTACACTTTCAGTA  
 AAAAAA

**FIGURE 92**

MALMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFFRGQFSS  
VVHLYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISQSYQK  
AIWELQVSALGSVPLISITGYVDRDIQLLCQSSGWFPPTAKWKGPQGQDLSTDSRTNRDMH  
GLFDVEISLTVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLF  
FGIVGLKIFFSKFQWKIQAELDWRKKGQAELRDARKHAVEVTLDPEAHKLCVSDLKTVT  
HRKAPQEVPHSEKRFRTRKSVVASQS FQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLS  
PDHGYWVLRNLNGEHLVFTLNPRFISVFPRTPTKIGVFLDYECGTISFFNINDQSLIYTLTC  
RFEGLLRPYIEYPSYNEQNGTPIVICPVTQESEKEASWQRASAI PETSNSESSSQATTPFLP  
RGEM

**Signal peptide:**

amino acids 1-17

**Transmembrane domain:**

amino acids 239-255



CGCATGGGTGCGCCCGGTGGCGGTGGCGGCGCGGTTGCGGAGGCTTCCTTGGTCGGATTGCA  
ACGAGGAGAAGATGACTGACCAACCGACTGGCTGAATGAATGAATGGCGGAGCCGAGCGCGC  
**CATG**AGGAGCCTGCCGAGCCTGGGCGGCCTCGCCCTGTTGTGCTGCGCCGCCGCCGCCGCCG  
CCGTGCGCTCAGCCGCCTCGGCGGGGAATGTCACCGGTGGCGGCGGGGCCGCGGGGCAGGTG  
GACGCGTCGCCGGGCCCCGGGTTCGCGGGGCGAGCCCAGCCACCCCTTCCCTAGGGCGACGGC  
TCCCACGGCCCAGGCCCCGAGGACCGGGCCCCCGCGCGCCACCGTCCACCGACCCCTGGCTG  
CGACTTCTCCAGCCCAGTCCCCGGAGACCACCCCTCTTTGGGCGACTGCTGGACCCTCTTCC  
ACCACCTTTTCAGGCGCCGCTCGGCCCTTCGCCGACCACCCCTCCGGCGGCGGAACGCACTTC  
GACCACCTCTCAGGCGCCGACCAGACCCGCGCCGACCACCCCTTTCGACGACCACTGGCCCGG  
CGCCGACCACCCCTGTAGCGACCACCGTACCGGCGCCACGACTCCCCGGACCCCGACCCCC  
GATCTCCCCAGCAGCAGCAACAGCAGCGTCCTCCCCACCCACCTGCCACCGAGGCCCCCTC  
TTCGCCTCCTCCAGAGTATGTATGTAACCTGCTCTGTGGTTGGAAGCCTGAATGTGAATCGCT  
GCAACCAGACCACAGGGCAGTGTGAGTGTGCGCCAGGTTATCAGGGGCTTCACTGTGAAACC  
TGCAAAGAGGGGCTTTTACCTAAATTACACTTCTGGGCTCTGTGAGCCATGTGACTGTAGTCC  
ACATGGAGCTCTCAGCATACCGTGCAACAGG**TAA**GCAACAGAGGGTGGAAGTGAAGTTTATT  
TTATTTTAGCAAGGGAAAAAAAAGGCTGCTACTCTCAAGGACCATACTGGTTTAAACAAAG  
GAGGATGAGGGTCATAGATTTACAAAATATTTTATATACTTTTATTCTCTTACTTTATATGT  
TATATTTAATGTCAGGATTTAAAAACATCTAATTTACTGATTTAGTTCTTCAAAGCACTAG  
AGTCGCCAATTTTTCTCTGGGATAATTTCTGTAAATTTTCATGGGAAAAAATTATTGAAGAAT  
AAATCTGCTTCTGGAAGGGCTTTCAGGCATGAAACCTGCTAGGAGGTTTAGAAATGTTCTT  
ATGTTTATTAATATAACATTGGAGTTTGAGGAAATTTGTTGTTTGGTTTATTTTTCTCTCTA  
ATCAAATTTCTACATTTGTTTCTTTGGACATCTAAAGCTTAACCTGGGGGTACCCTAATTTA  
TTTAACTAGTGGAAGTAGACTGGTTTTACTCTATTTACCAGTACATTTTTTGAGACCAAAG  
TAGATTAAGCAGGAATTATCTTTAAACTATTATGTTATTTGGAGGTAATTTAATCTAGTGGA  
ATAATGTACTGTTATCTAAGCATTTCCTTGTACTGCACTGAAAGTAATTATTCTTTGACCT  
TATGTGAGGCACATTGGCTTTTTGTGGACCCCAAGTCAAAAACCTGAAGAGACAGTATTAAT  
AATGAAAAAAATAATGACAGGTTATACTAGTGAACCTTGGGTATAACCCAAGATCTGCTGC  
CACTTACGAGCTGTGTTCTTTGGGCAAGTAATTTCTTTCACTGAGCTTGTTTCTTCTCAAG  
GTTGTTGTGAAGATTAAATGAGTTGATATATATAAAATGCCTAGCACATGTCACCTCAATAAA  
TTCTGGTTTGTTTTTAATTTCAAAGGAATATTATGGACTGAAATGAGAGAACATGTTTTAAGA  
ACTTTTAGCTCCTTGACAAAGAAGTGCTTTATACTTTAGCACTAAATATTTTAAATGCTTTA  
TAAATGATATTATACTGTTATGGAATATTGTATCATATTGTAGTTTATTAAAAATGTAGAAG  
AGGCTGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTTTGGGAGGCCAAGGCGGGTGGAT  
CACTTGAGGCCAGGAGTTCTAGATGAGCCTGGCCAGCACAGTGAAACCCCGTCTCTACTAAA  
AATACAAACAAATTAGCTGGGCGTGGTGGCACACACCTGTAGTCCAGCTACTCGGGAGGCT  
GAGGCAGGAGAATCGGTTGAACCCGGGAGGTGGAGGTTGCAGTGAGCTGAGATCGCGCCACT  
GCACTCCAGCCTGGTGAGAGAGGGGAGACTCTGTCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 94**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64952

><subunit 1 of 1, 258 aa, 1 stop

><MW: 25716, pI: 8.13, NX(S/T): 5

MRSLSLGLGGLALLCCAAAAA AVASAASAGNVTGGGGAAGQVDASPGPGLRGEP SHPFPRATA  
PTAQAPRTGPPRATVHRPLAATSPAQSPETTPLWATAGPSSTTFQAPLGPSPTTPPAAERTS  
TTSQAPTRPAPTTLSTTTGAPPTPVATTVPAPTTPRTPTPDLPSSSNSSVLPTPPATEAPS  
SPPPEYVCNCSVVGSLNVNRCNQTTGQCECRPGYQGLHCETCKEGFYLNYSGLCQPCDCSP  
HGALSIPCNR

**Important features of the protein:**

**Signal peptide:**

amino acids 1-25

**N-glycosylation sites.**

amino acids 30-33, 172-175, 195-198, 208-211, 235-238

**EGF-like domain cysteine pattern signature.**

amino acids 214-226.

**FIGURE 95**

TGCGGCGCAGTGTAGACCTGGGAGGATGGGCGGCCTGCTGCTGGCTGCTTTTCTGGCTTTGG  
TCTCGGTGCCCAGGGCCCAGGCCGTGTGGTTGGGAAGACTGGACCCTGAGCAGCTTCTTGGG  
CCCTGGTACGTGCTTGCGGTGGCCTCCCGGGAAAAGGGCTTTGCCATGGAGAAGGACATGAA  
GAACGTCGTGGGGGTGGTGGTGACCCTCACTCCAGAAAACAACCTGCGGACGCTGTCCTCTC  
AGCACGGGCTGGGAGGGTGTGACCAGAGTGTGATGGACCTGATAAAGCGAAACTCCGGATGG  
GTGTTTGAGAATCCCTCAATAGGCGTGCTGGAGCTCTGGGTGCTGGCCACCAACTTCAGAGA  
CTATGCCATCATCTTCACTCAGCTGGAGTTCGGGGACGAGCCCTTCAACACCGTGGAGCTGT  
ACAGTCTGACGGAGACAGCCAGCCAGGAGGCCATGGGGCTCTTCACCAAGTGGAGCAGGAGC  
CTGGGCTTCCTGTCACAGTAGCAGGCCCCAGCTGCAGAAGGACCTCACCTGTGCTCACAAGAT  
CCTTCTGTGAGTGTGCGTCCCCAGTAGGGATGGCGCCACAGGGTCCTGTGACCTCGGCCA  
GTGTCCACCCACCTCGCTCAGCGGCTCCCGGGGGCCAGCACCAGCTCAGAATAAAGCGATTC  
CACAGCA

TCTCGGTGCCCAGGGCCCAGGCCGTGTGGTTGGGAAGACTGGACCCTGAGCAGCTTCTTGGG

## **FIGURE 96**

MGGLLLAAFLALVSVPRQA VWLGRLDPEQLLGPWYVLAVASREKGFAMEKDMKNVVGVVVT  
LTPENNLRTLSSQHGLGGCDQSVMDLIKRN SGWVFENPSIGVLELWVLATNFRDYAIIFTQL  
EFGDEPFNTVELYSLTETASQEAMGLFTKWSRSLGFLSQ

**Signal peptide:**

amino acids 1-20

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 97**

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGAC**ATG**CTGCTGCTGCTGCTGCC  
CCTGCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTT  
CCGTGACGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCAT  
GGCTGGATTTACCCCTGGCCCAGTAGTTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGA  
CCAGGATGCTCCAGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACC  
GATTCCACCTCCTTGGGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGA  
AGAAGTGATGCGGGGAGATACTTCTTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAA  
ACATCACCGGCTCTCTGTGAATGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAG  
GCACCCTGGAGTCCGGCTGCCCCCAGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAG  
GGGACACCCCCTATGATCTCCTGGATAGGGACCTCCGTGTCCCCCTGGACCCCTCCACCAC  
CCGCTCCTCGGTGCTCACCCCTCATCCACAGCCCCAGGACCATGGCACCAGCCTCACCTGTC  
AGGTGACCTTCCCTGGGGCCAGCGTGACCACGAACAAGACCGTCCATCTCAACGTGTCCTAC  
CCGCCTCAGAACTTGACCATGACTGTCTTCCAAGGAGACGGCACAGTATCCACAGTCTTGGG  
AAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTCTCTGCGCCTGGTCTGTGCAGTTGATG  
CAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGAGAGGCCTGACCCTGTGCCCC  
TCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCACCTGAGGGATGCAGCTGA  
ATTCACCTGCAGAGCTCAGAACCCTCTCGGCTCTCAGCAGGTCTACCTGAACGTCTCCCTGC  
AGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGGAGCTGGAGCCACAGCCCTG  
GTCTTCCTGTCCTTCTGCGTCATCTTCGTTGTAGTGAGGTCCTGCAGGAAGAAATCGGCAAG  
GCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTTCAGCCT  
CTCAGGGGCCCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCCAGCT  
TCTGCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGT  
GAAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCC  
ACAGAT**TGA**GAAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCCTCCAGGCAAGGGA  
GAAGTCAGAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACT  
ATGAATTATGTGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAAT  
CCACACTGTGCCCTCCCTTTTATTTTTTTAACTAAAAGACAGACAAATTCCTA

**FIGURE 98**

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSPSHGWIYPGPVVHGYWF  
REGANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKG  
SIKWNKHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVS  
PLDPSTTRSSVLTLLIPQPQDHGTSLTCQVTFPGASVTTNKTVHLNVSYPPQNLMTVFQGDG  
TVSTVLGNGSSLSLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWV  
HLRDAAEFTCRAQNPLGSQQVYLNVSLSQKATSGVTQGVVGGAGATALVFLSFCVIFVVVRS  
CRKKSARPAAGVGDGIEDANAVRGSSASQGPLEPWAEDSPPDQPPPASARSSVGEGELQYA  
SLSFQMKPWDSRGQEATDTEYSEIKIHR

**Signal peptide:**

amino acids 1-15

**Transmembrane domain:**

amino acids 351-370

GACGCCCCAGTGACCTGCCGAGGTCGGCAGCACAGAGCTCTGGAG**ATGA**AAGACCCTGTTCCTG  
GGTGTACGCTCGGCCTGGCCGCTGCCCTGTCTTCACCCTGGAGGAGGAGGATATCACAGG  
GACCTGGTACGTGAAGGCCATGGTGGTCGATAAGGACTTTCGGAGGACAGGAGGCCAGGA  
AGGTGTCCCCAGTGAAGGTGACAGCCCTGGGCGGTGGGAAGTTGGAAGCCACGTTACCTTC  
ATGAGGGAGGATCGGTGCATCCAGAAGAAAATCCTGATGCGGAAGACGGAGGAGCCTGGCAA  
ATACAGCGCCTATGGGGGCAGGAAGCTCATGTACCTGCAGGAGCTGCCCAGGAGGGACCACT  
ACATCTTTTACTGCAAAGACCAGCACCATGGGGGCCTGCTCCACATGGGAAGCTTGTGGGT  
AGGAATTCTGATACCAACCGGGAGGCCCTGGAAGAATTTAAGAAATTGGTGCAGCGCAAGGG  
ACTCTCGGAGGAGGACATTTTCACGCCCTGCAGACGGGAAGCTGCGTTCCCGAACAC**TAGG**  
CAGCCCCCGGGTCTGCACCTCCAGAGCCACCCTACCACCAGACACAGAGCCCGGACCACCT  
GGACCTACCCTCCAGCCATGACCCTTCCTTGCTCCCACCCACCTGACTCCAAATAAAGTCCT  
TTTCCCCCAA

**FIGURE 100**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65404

<subunit 1 of 1, 170 aa, 1 stop

<MW: 19457, pI: 9.10, NX(S/T): 0

MKTLFLGVTLGLAAALSFTLEEEDITGTWYVKAMVVDKDFPEDRRPRKVSPVKVTALGGGKL  
EATFTFMREDRCIQKKILMRKTEEPGKYSAYGGRKLMYLQELPRRDHYIFYCKDQHHGGLLH  
MGKLVGRNSDTNREALEEFKKLVQRKGLSEEDIFTPLQTGSCVPEH

**Important features:****Signal peptide:**

amino acids 1-17

100/249



**FIGURE 101**

GTTCCGCAGATGCAGAGGTTGAGGTGGCTGCGGGACTGGAAGTCATCGGGCAGAGGTCTCAC  
AGCAGCCAAGGAACCTGGGGCCCGCTCCTCCCCCTCCAGGCC**ATG**AGGATTCTGCAGTTAA  
TCCTGCTTGCTCTGGCAACAGGGCTTGTAGGGGGAGAGACCAGGATCATCAAGGGGTTCGAG  
TGCAAGCCTCACTCCCAGCCCTGGCAGGCAGCCCTGTTTCGAGAAGACGCGGCTACTCTGTGG  
GGCGACGCTCATCGCCCCAGATGGCTCCTGACAGCAGCCCACTGCCTCAAGCCCCGCTACA  
TAGTTCACCTGGGGCAGCACAACTCCAGAAGGAGGAGGGCTGTGAGCAGACCCGGACAGCC  
ACTGAGTCCTTCCCCACCCCGGCTTCAACAACAGCCTCCCCAACAAAGACCACCGCAATGA  
CATCATGCTGGTGAAGATGGCATCGCCAGTCTCCATCACCTGGGCTGTGCGACCCCTCACCC  
TCTCCTCACGCTGTGTCACTGCTGGCACCAGCTGCCTCATTTCCGGCTGGGGCAGCACGTCC  
AGCCCCCAGTTACGCCTGCCTCACACCTTGCGATGCGCCAACATCACCATCATTGAGCACCA  
GAAGTGTGAGAACGCCTACCCCGGCAACATCACAGACACCATGGTGTGTGCCAGCGTGCAGG  
AAGGGGGCAAGGACTCCTGCCAGGGTGACTCCGGGGGCCCTCTGGTCTGTAACCAGTCTCTT  
CAAGGCATTATCTCCTGGGGCCAGGATCCGTGTGCGATCACCCGAAAGCCTGGTGTCTACAC  
GAAAGTCTGCAAATATGTGGACTGGATCCAGGAGACGATGAAGAACAAT**TAG**ACTGGACCCA  
CCCACCACAGCCCATCACCCCTCCATTTCCACTTGGTGTGTTGGTTCCTGTTCACTCTGTTAAT  
AAGAAACCCTAAGCCAAGACCCTCTACGAACATTCTTTGGGCCTCCTGGACTACAGGAGATG  
CTGTCACTTAATAATCAACCTGGGGTTCGAAATCAGTGAGACCTGGATTCAAATTCTGCCTT  
GAAATATTGTGACTCTGGGAATGACAACACCTGGTTTGTTCTCTGTTGTATCCCCAGCCCCA  
AAGACAGCTCCTGGCCATATATCAAGGTTTCAATAAATATTTGCTAAATGAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 102**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65405

<subunit 1 of 1, 250 aa, 1 stop

<MW: 27466, pI: 8.87, NX(S/T): 4

MRILQLILLALATGLVGGETRIIKGFECKPHSQPWQAALFEKTRLLCGATLIAPRWLLTAAH  
 CLKPRYIVHLGQHNLQKEEGCEQTRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPVSITW  
 AVRPLTLSSRCVTAGTSCSLISGWGSTSSPQLRLPHTLRCANITIIIEHQKCENAYPGNITDTM  
 VCASVQEGGKDSCQGDSSGGLVCNQSLQGIISWGQDPCAITRKPGVYTKVCKYVDWIQETMKNN

**Important features:**

**Signal peptide:**

amino acids 1-18

**Serine proteases, trypsin family, histidine active site.**

amino acids 58-63

**N-glycosylation sites.**

amino acids 99-102, 165-168, 181-184, 210-213

**Glycosaminoglycan attachment site.**

amino acids 145-148

**Kringle domain proteins.**

amino acids 197-209, 47-64

**Serine proteases, trypsin family, histidine protein**

amino acids 199-209, 47-63, 220-243

**Apple domain proteins**

amino acids 222-249, 189-222

**FIGURE 103**

GAGCAGTGTTCTGCTGGAGCCG**ATG**CCAAAAACCATGCATTTCTTATTCAGATTCATTGTTT  
TCTTTTATCTGTGGGGCCTTTTTACTGCTCAGAGACAAAAGAAAGAGGAGAGCACCGAAGAA  
GTGAAAATAGAAGTTTTGCATCGTCCAGAAAAGTCTCTAAGACAAGCAAGAAGGGAGACCT  
ACTAAATGCCCATTATGACGGCTACCTGGCTAAAGACGGCTCGAAATTCTACTGCAGCCGGA  
CACAAAATGAAGGCCACCCCAAATGGTTTGTTCCTTGGTGTGGGCAAGTCATAAAAGGCCTA  
GACATTGCTATGACAGATATGTGCCCTGGAGAAAAGCGAAAAGTAGTTATACCCCTTCATT  
TGCATACGGAAAGGAAGGCTATGCAGAAGGCAAGATTCCACCGGATGCTACATTGATTTTTG  
AGATTGAACTTTATGCTGTGACCAAAGGACCACGGAGCATTGAGACATTTAAACAAATAGAC  
ATGGACAATGACAGGCAGCTCTCTAAAGCCGAGATAAACCTCTACTTGCAAAGGGAATTTGA  
AAAAGATGAGAAGCCACGTGACAAGTCATATCAGGATGCAGTTTTAGAAAGATATTTTTAAGA  
AGAATGACCATGATGGTGATGGCTTCATTTCTCCAAGGAATACAATGTATACCAACACGAT  
GAACTA**TAG**CATATTTGTATTTCTACTTTTTTTTTTTTAGCTATTTACTGTACTTTATGTATA  
AAACAAAGTCACTTTTCTCCAAGTTGTATTTGCTATTTTCCCCTATGAGAAGATATTTTGA  
TCTCCCCAATACATTGATTTTGGTATAATAAATGTGAGGCTGTTTTGCAAACCTTAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 104**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65406

<subunit 1 of 1, 222 aa, 1 stop

<MW: 25794, pI: 6.24, NX(S/T): 1

MPKTMHFLFRFIVFFYLWGLFTAQRQKKEESTEEVKIEVLHRPENCSKTSKKGDLLNAHYDG  
YLAKDGSKFYCSRTQNEGHPKWFVLGVGQVIKGLDIAMTDMCPGEKRVVIPP SFAYGKEGY  
AEGKIPPDATLIFEIELYAVTKGPRSIETFKQIDMDNDRQLSKAEINLYLQREFEKDEKPRD  
KSYQDAVLEDIFKKNDHGDGFISPKEYNVYQHDEL

**Important features:**

**Endoplasmic reticulum targeting sequence.**

amino acids 219-222

**N-glycosylation site.**

amino acids 45-48

**FKBP-type peptidyl-prolyl cis-trans isomerase**

amino acids 87-223, 129-142

**EF-hand calcium-binding domain proteins**

amino acids 202-214, 195-214

**FIGURE 105**

CAGAA**ATG**CAGGGACCATTGCTTCTTCCAGGCCTCTGCTTTCTGCTGAGCCTCTTTGGAGCT  
GTGACTCAGAAAACCAAACTTCCTGTGCTAAGTGCCCCCAAATGCTTCCTGTGTCAATAA  
CACTCACTGCACCTGCAACCATGGATATACTTCTGGATCTGGGCAGAACTATTCACATTCC  
CCTTGGAGACATGTAACGCCAGGCATGGTGGCTCGCGCCTG**TAA**TCCCAGTTCTTTGGGAAG  
CCAAGGCAGGTGGATCACCTGAGGTCAGGAGTTTGAGACCAGCCTGGCCAACATAGTGAAAC  
CCCGTGTCTACTAAAAATACAAAAATCAGCCGGGCGTGGTGGTGCATGCCTGCAATCCCAGT  
TACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCTCAGGAGGCAGAAGTTGCAGTGAACCC  
AGATCCTGCCATTGCACTCCAGCATGGATGACAGAGCAAGACTCCGTCTCAAAAAGAAAAGA  
TAGTTTCTTGTTTCATTTTCGCGACTGCCCTCTCAGTGTTTCCTGGGATCCCCTCCCAAATAA  
AGTACTTATATTCTC

106/249

## **FIGURE 106**

MQGPLLLPGLCFLLSLFGAVTQKTKTSCAKCPPNASCVNNTHTCTCNHGYTSGSGQKLFTFPL  
ETCNARHGGSRL

**Signal peptide:**

amino acids 1-18

106/249

**FIGURE 107**

CAAGCAGGTCATCCCCTTGGTGACCTTCAAAGAGAAGCAGAGAGGGCAGAGGTGGGGGGGCAC  
AGGGAAAGGGTGACCTCTGAGATTCCCCTTTTCCCCCAGACTTTGGAAGTGACCCACC**ATGG**  
GGCTCAGCATCTTTTTGCTCCTGTGTGTTCTTGGGCTCAGCCAGGCAGCCACACCGAAGATT  
TTCAATGGCACTGAGTGTGGGCGTAACTCACAGCCGTGGCAGGTGGGGCTGTTTGAGGGGCAC  
CAGCCTGCGCTGCGGGGGTGTCCTTATTGACCACAGGTGGGTCCCTCACAGCGGCTCACTGCA  
GCGGCAGCAGGTACTGGGTGCGCCTGGGGGAACACAGCCTCAGCCAGCTCGACTGGACCGAG  
CAGATCCGGGCACAGCGGCTTCTCTGTGACCCATCCCGGCTACCTGGGAGCCTCGACGAGCCA  
CGAGCACGACCTCCGGCTGCTGCGGCTGCGCCTGCCCGTCCGCGTAACCAGCAGCGTTCAAC  
CCCTGCCCCCTGCCCAATGACTGTGCAACCGCTGGCACCGAGTGCCACGTCTCAGGCTGGGGC  
ATCACCAACCACCCACGGAACCCATTCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATCGT  
CTCCCATGCCACCTGCCATGGTGTGTATCCCGGGAGAATCACGAGCAACATGGTGTGTGCAG  
GCGGCGTCCCGGGGCAGGATGCCTGCCAGGGTGATTCTGGGGGCCCCCTGGTGTGTGGGGGA  
GTCCTTCAAGGTCTGGTGTCTTGGGGGTCTGTGGGGCCCTGTGGACAAGATGGCATCCCTGG  
AGTCTACACCTATATTTGCAAGTATGTGGACTGGATCCGGATGATCATGAGGAACAAC**TGAC**  
CTGTTTTCCTCCACCTCCACCCCCACCCCTTAACCTGGGTACCCCTCTGGCCCTCAGAGCACC  
AATATCTCCTCCATCACTTCCCCTAGCTCCACTCTTGTTGGCCTGGGAACTTCTTGGAACCTT  
TAACTCCTGCCAGCCCTTCTAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGA  
ATAAATATAAATGAAGGAGGGGCAAAAAAAAAAAAA

## **FIGURE 108**

MGLSIFLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWVLTAAH  
CSGSRYWVRLGEHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLRLPVRVTSSV  
QPLPLPNDCATAGTECHVSGWGITNHPRNPFDPDLLQCLNLSIVSHATCHGVYPGRITSNMVC  
AGGVPGQDACQGDSGGPLVCGGVLQGLVSWGSGVGPCGQDGIPGVYTYICKYVDWIRMIMRNN

**Signal peptide:**

amino acids 1-17

108/249



**FIGURE 109**

GCGGCCACACGCAGCTAGCCGGAGCCCGGACCAGGCGCCTGTGCCTCCTCCTCGTCCCTCGC  
 CGCGTCCGCGAAGCCTGGAGCCGGCGGGAGCCCCGCGCTCGCC**ATG**TCGGGCGAGCTCAGCA  
 ACAGGTTCCAAGGAGGGAAGGCGTTTCGGCTTGCTCAAAGCCCGGCAGGAGAGGAGGCTGGCC  
 GAGATCAACCGGGAGTTTCTGTGTGACCAGAAGTACAGTGATGAAGAGAACCTTCCAGAAAA  
 GCTCACAGCCTTCAAAGAGAAGTACATGGAGTTTGACCTGAACAATGAAGGCGAGATTGACC  
 TGATGTCTTTAAAGAGGATGATGGAGAAGCTTGGTGTCCCCAAGACCCACCTGGAGATGAAG  
 AAGATGATCTCAGAGGTGACAGGAGGGGTGAGTACACTATATCCTACCGAGACTTTGTGAA  
 CATGATGCTGGGGAAACGGTTCGGCTGTCCTCAAGTTAGTCATGATGTTTGAAGGAAAAGCCA  
 ACGAGAGCAGCCCCAAGCCAGTTGGCCCCCCTCCAGAGAGAGACATTGCTAGCCTGCCCT**TGA**  
 GGACCCCGCCTGGACTCCCCAGCCTTCCCACCCCATACCTCCCTCCCGATCTTGCTGCCCTT  
 CTTGACACACTGTGATCTCTCTCTCTCATTTGTTTGGTCATTGAGGGTTTGTGTTGTGTTT  
 TCATCAATGTCTTTGTAAAGCACAAATTATCTGCCTTAAAGGGGCTCTGGGTTCGGGGAATCC  
 TGAGCCTTGGGTCCCCCTCCTCTCTTCTTCCCTCCTTCCCCGCTCCCTGTGCAGAAGGGCTG  
 ATATCAAACCAAAAAGTAGAGGGGGCAGGGCCAGGGCAGGGAGGCTTCCAGCCTGTGTTCCC  
 CTCACCTGGAGGAACCAGCACTCTCCATCCTTTTCAAGAAAGTCTCCAAGCCAAGTTCAGGCTC  
 ACTGACCTGGCTCTGACGAGGACCCCAGGCCACTCTGAGAAGACCTTGGAGTAGGGACAAGG  
 CTGCAGGGCCTCTTTCGGGTTTTCCTTGGACAGTGCCATGGTTCCAGTGCTCTGGTGTACCC  
 AGGACACAGCCACTCGGGGCCCCGCTGCCCCAGCTGATCCCCACTCATTCCACACCTCTTCT  
 CATCCTCAGTGATGTGAAGGTGGGAAGGAAAGGAGCTTGGCATTGGGAGCCCTTCAAGAAGG  
 TACCAGAAGGAACCCTCCAGTCCTGCTCTCTGGCCACACCTGTGCAGGCAGCTGAGAGGCAG  
 CGTGCAGCCCTACTGTCCCTTACTGGGGCAGCAGAGGGCTTCGGAGGCAGAAAGTGAGGCCTG  
 GGGTTTGGGGGGAAAGGTCAGCTCAGTGCTGTTCCACCTTTTAGGGAGGATACTGAGGGGAC  
 CAGGATGGGAGAATGAGGAGTAAAATGCTCACGGCAAAGTCAGCAGCACTGGTAAGCCAAGA  
 CTGAGAAATACAAGGTTGCTTGTCTGACCCCAATCTGCTTGAAAAAAAAAAAAAAAAAAAA

Figure 1 consists of 12 horizontal bar charts, labeled (a) through (l), arranged vertically. Each chart displays the percentage of total protein for 12 different protein types (A through L) across 12 different conditions (1 through 12). The y-axis for each chart represents the percentage of total protein, ranging from 0 to 100. The x-axis represents the protein type. The bars are color-coded: A (black), B (white), C (grey), D (dark grey), E (light grey), F (medium grey), G (dark grey), H (light grey), I (medium grey), J (dark grey), K (light grey), L (medium grey). The data shows varying distributions of protein types across the conditions, with some types being more prevalent in certain conditions than others.

Figure 1 consists of 12 horizontal bar charts, labeled (a) through (l), arranged vertically. Each chart displays the percentage of total protein for 12 different protein types (A through L) across 12 different conditions (1 through 12). The y-axis for each chart represents the percentage of total protein, ranging from 0 to 100. The x-axis represents the protein type. The bars are color-coded: A (black), B (white), C (grey), D (dark grey), E (light grey), F (medium grey), G (dark grey), H (light grey), I (medium grey), J (dark grey), K (light grey), L (medium grey). The data shows varying distributions of protein types across the conditions, with some types being more prevalent in certain conditions than others.

**FIGURE 111A**

CGCGCTCCCCGCGCGCCTCCTCGGGCTCCACGCGTCTTGCCCCGCAGAGGCAGCCTCCTCCA  
GGAGCGGGGCCCTGCACACC**ATG**GGCCCCGGGTGGGCAGGGGTGGCGCCGCCGTGCGCGCC  
CGCCTGGCGCTGGCCTTGGCGCTGGCGAGCGTCTTGAGTGGGCCTCCAGCCGTGCGCTGCCC  
CACCAAGTGTACCTGCTCCGCTGCCAGCGTGGACTGCCACGGGCTGGGCCTCCGCGCGGTTC  
CTCGGGGCATCCCCCGCAACGCTGAGCGCCTTGACCTGGACAGAAATAATATCACCAGGATC  
ACCAAGATGGACTTCGCTGGGCTCAAGAACCTCCGAGTCTTGATCTGGAAGACAACCAGGT  
CAGCGTCATCGAGAGAGGGCGCCTTCCAGGACCTGAAGCAGCTAGAGCGACTGCGCCTGAACA  
AGAATAAGCTGCAAGTCTTCCAGAATTGCTTTTCCAGAGCACGCCGAAGCTCACCAGACTA  
GATTTGAGTGAAAACCAGATCCAGGGGATCCCGAGGAAGGCGTTCCGCGGCATCACCAGTGT  
GAAGAACCTGCAACTGGACAACAACCACATCAGCTGCATTGAAGATGGAGCCTTCCGAGCGC  
TGCGCGATTTGGAGATCCTTACCCTCAACAACAACAACATCAGTCGCATCCTGGTCACCAGC  
TTCAACCACATGCCGAAGATCCGAACCTCTGCGCCTCCACTCCAACCACCTCTACTGCGACTG  
CCACCTGGCCTGGCTCTCGGATTGGCTGCGACAGCGACGGACAGTTGGCCAGTTCACACTCT  
GCATGGCTCCTGTGCATTTGAGGGGCTTCAACGTGGCGGATGTGCAGAAGAAGGAGTACGTG  
TGCCAGCCCCCACTCGGAGCCCCCATCCTGCAATGCCAACTCCATCTCCTGCCCTTCGCC  
CTGCACGTGCAGCAATAACATCGTGGACTGTCGAGGAAAGGGCTTGATGGAGATTCCTGCCA  
ACTTGCCGGAGGGCATCGTCGAAATACGCCTAGAACAGAACTCCATCAAAGCCATCCCTGCA  
GGAGCCTTCACCCAGTACAAGAACTGAAGCGAATAGACATCAGCAAGAATCAGATATCGGA  
TATTGCTCCAGATGCCTTCCAGGGCCTGAAATCACTCACATCGCTGGTCTGTATGGGAACA  
AGATCACCGAGATTGCCAAGGGACTGTTTGATGGGCTGGTGTCCCTACAGCTGCTCCTCCTC  
AATGCCAACAAGATCAACTGCCTGCGGGTGAACACGTTTCAGGACCTGCAGAACCTCAACTT  
GCTCTCCCTGTATGACAACAAGCTGCAGACCATCAGCAAGGGGCTCTTCGCCCCCTCTGCAGT  
CCATCCAGACACTCCACTTAGCCCCAAAACCCATTTGTGTGCGACTGCCACTTGAAGTGGCTG  
GCCGACTACCTCCAGGACAACCCCATCGAGACAAGCGGGGCCCCGCTGCAGCAGCCCCGCGCCG  
ACTCGCCAACAAGCGCATCAGCCAGATCAAGAGCAAGAAGTTCCGCTGCTCAGGCTCCGAGG  
ATTACCGCAGCAGGTTTCAGCAGCGAGTGCTTCATGGACCTCGTGTGCCCCGAGAAGTGTGCG  
TGTGAGGGCACGATTGTGGACTGCTCCAACCAGAAGCTGGTCCGCATCCCAAGCCACCTCCC  
TGAATATGTCACCGACCTGCGACTGAATGACAATGAGGTATCTGTTCTGGAGGGCACTGGCA  
TCTTCAAGAAGTTGCCAACCTGCGGAAAATAAATCTGAGTAACAATAAGATCAAGGAGGTG  
CGAGAGGGAGCTTTCGATGGAGCAGCCAGCGTGCAGGAGCTGATGCTGACAGGGAACCAGCT  
GGAGACCGTGCACGGGCGCGTGTTCGCTGGCCTCAGTGGCCTCAAAACCTTGATGCTGAGGA  
GTAACCTTGATCAGCTGTGTGAGTAATGACACCTTTGCCGGCCTGAGTTCGGTGAGACTGCTG  
TCCCTCTATGACAATCGGATCACCACCATCACCCCTGGGGCCTTCACCACGCTTGTCTCCCT  
GTCCACCATAAACCTCCTGTCCAACCCCTTCAACTGCAACTGCCACCTGGCCTGGCTCGGCA  
AGTGTTGAGGAAGAGGGCGGATCGTCAGTGGGAACCCTAGGTGCCAGAAGCCATTTTTCCTC  
AAGGAGATTCCCATCCAGGATGTGGCCATCCAGGACTTCACCTGTGATGGCAACGAGGAGAG  
TAGCTGCCAGCTGAGCCCCGCGCTGCCCGGAGCAGTGCACCTGTATGGAGACAGTGGTGCAT  
GCAGCAACAAGGGGCTCCGCGCCCTCCCCAGAGGCATGCCAAGGATGTGACCGAGCTGTAC  
CTGGAAGGAAACCACCTAACAGCCGTGCCAGAGAGCTGTCCGCCCTCCGACACCTGACGCT  
TATTGACCTGAGCAACAACAGCATCAGCATGCTGACCAATTACACCTTCAGTAACATGTCTC  
ACCTCTCCACTCTGATCCTGAGCTACAACCGGCTGAGGTGCATCCCCGTCCACGCCTTCAAC  
GGGCTGCGGTCCCTGCGAGTGCTAACCCCTCCATGGCAATGACATTTCCAGCGTTTCTGAAGG  
CTCCTTCAACGACCTCACATCTCTTCCCATCTGGCGCTGGGAACCAACCCACTCCACTGTG  
ACTGCAGTCTTCGGTGGCTGTGCGAGTGGGTGAAGGCGGGGTACAAGGAGCCTGGCATCGCC  
CGCTGCAGTAGCCCTGAGCCCATGGCTGACAGGCTCCTGCTCACCACCCCAACCCACCGCTT  
CCAGTGCAAAGGGCCAGTGGACATCAACATTGTGGCCAAATGCAATGCCTGCCTCTCCAGCC  
CGTGCAAGAATAACGGGACATGCACCCAGGACCCTGTGGAGCTGTACCGCTGTGCCTGCCCC

**FIGURE 111B**

TACAGCTACAAGGGCAAGGACTGCACTGTGCCCATCAACACCTGCATCCAGAACCCCTGTCA  
 GCATGGAGGCACCTGCCACCTGAGTGACAGCCACAAGGATGGGTTTCAGCTGCTCCTGCCCTC  
 TGGGCTTTGAGGGGCAGCGGTGTGAGATCAACCCAGATGACTGTGAGGACAACGACTGCGAA  
 AACAATGCCACCTGCGTGGACGGGATCAACAACCTACGTGTGTATCTGTCCGCCTAACTACAC  
 AGGTGAGCTATGCGACGAGGTGATTGACCACTGTGTGCCTGAGCTGAACCTCTGTCAGCATG  
 AGGCCAAGTGTCATCCCCCTGGACAAAGGATTGAGCTGCGAGTGTGTCCCTGGCTACAGCGGG  
 AAGCTCTGTGAGACAGACAATGATGACTGTGTGGCCCAAGTGCCGCCACGGGGCCCAGTG  
 CGTGGACACAATCAATGGCTACACATGCACCTGCCCCCAGGGCTTCAGTGGACCCTTCTGTG  
 AACACCCCCCACCCATGGTCCTACTGCAGACCAGCCCATGCGACCAGTACGAGTGCCAGAAC  
 GGGGCCCAGTGTCATCGTGGTGCAGCAGGAGCCACCTGCCGCTGCCACCAGGCTTCGCCGG  
 CCCCAGATGCGAGAAGCTCATCACTGTCAACTTCGTGGGCAAAGACTCCTACGTGGAACCTGG  
 CCTCCGCCAAGGTCCGACCCCAGGCCAACATCTCCCTGCAGGTGGCCACTGACAAGGACAAC  
 GGCATCCTTCTCTACAAAGGAGACAATGACCCCCCTGGCACTGGAGCTGTACCAGGGCCACGT  
 GCGGCTGGTCTATGACAGCCTGAGTTCCCCTCCAACCACAGTGTACAGTGTGGAGACAGTGA  
 ATGATGGGCAGTTTCACAGTGTGGAGCTGGTGACGCTAAACCAGACCCTGAACCTAGTAGTG  
 GACAAAGGAACTCCAAAGAGCCTGGGGAAGCTCCAGAAGCAGCCAGCAGTGGGCATCAACAG  
 CCCCCTCTACCTTGGAGGCATCCCCACCTCCACCGGCCTCTCCGCCTTGCGCCAGGGCACGG  
 ACCGGCCTCTAGGCGGGCTTCCACGGATGCATCCATGAGGTGCGCATCAACAACGAGCTGCAG  
 GACTTCAAGGCCCTCCCACCACAGTCCCTGGGGGTGTCACCAGGCTGCAAGTCTGCAACCGT  
 GTGCAAGCACGGCCTGTGCCGCTCCGTGGAGAAGGACAGCGTGGTGTGCGAGTGCCGCCAG  
 GCTGGACCGGCCCCACTCTGCGACCAGGAGGCCCGGGACCCCTGCCTCGGCCACAGATGCCAC  
 CATGGAAAATGTGTGGCAACTGGGACCTCATAATGTGCAAGTGTGCCGAGGGCTATGGAGG  
 GGACTTGTGTGACAACAAGAATGACTCTGCCAATGCCTGCTCAGCCTTCAAGTGTCAACCATG  
 GGCAGTGCCACATCTCAGACCAAGGGGAGCCCTACTGCCTGTGCCAGCCCGGCTTTAGCGGC  
 GAGCACTGCCAACAAGAGAATCCGTGCCTGGGACAAGTAGTCCGAGAGGTGATCCGCCGCCA  
 GAAAGGTATGCATCATGTGCCACAGCCTCCAAGGTGCCCATCATGGAATGTCGTGGGGGCT  
 GTGGGCCCCAGTGCTGCCAGCCCACCCGCAGCAAGCGGCGGAAATACGTCTTCCAGTGCACG  
 GACGGCTCCTCGTTTGTAGAAGAGGTGGAGAGACACTTAGAGTGCGGCTGCCTCGCGTGTTC  
**CTAA**GCCCCCTGCCCGCCTGCCTGCCACCTCTCGGACTCCAGCTTGATGGAGTTGGGACAGCC  
 ATGTGGGACCCCTGGTGATTCAGCATGAAGGAAATGAAGCTGGAGAGGAAGGTAAAGAAGA  
 AGAGAATATTAAGTATATTGTAAATAAACAACAAAAATAGAACTTAAAAAAAAAAAAAAAAAA  
 AAAAAA

**FIGURE 112**

MAPGWAGVGAAVRARLALALASVLSGPPAVACPTKCTCSAASVDCHGLGLRAVPRGIPRN  
 AERLDLDRNNITRITKMDFAGLKNLRVLHLEDNQVSVIERGAFQDLKQLERLRLNKNKLQVL  
 PELLFQSTPKLTRLDLSENQIQGIPRKAFRGITDVKNLQLDNNHISCIEDGAFRALRDLEIL  
 TLNNNNISRILVTSFNHMPKIRTLRLHSNHLCDCHLAWLSDWLRQRRTVGQFTLCMAPVHL  
 RGFNVADVQKEYVCPAPHSEPPSCNANSISCPSPCTCSNNIVDCRGKGLMEIPANLPEGIV  
 EIRLEQNSIKAIPAGAFTQYKKLKRIDISKNQISDIAPDAFQGLKSLTSLVLYGNKITEIAK  
 GLFDGLVSLQLLLLNANKINCLRVNTFQDLQNLNLLSLYDNKLQTISKGLFAPLQSIQTLHL  
 AQNPFFVCDCHLKWLADYLQDNPIETSGARCSSPRRLANKRISQIKSKKFRCSGSEDIRSRFS  
 SECFMDLVCPEKCRCEGTIVDCSNQKLVRIPSHLPEYVTDLRLNDNEVSVLEATGIFKKLPN  
 LRKINLSNNKIKEVREGAFDGAASVQELMLTGNQLETVHGRVFRGLSGLKTLMLRSNLISCV  
 SNDTFAGLSSVRLLSLYDNRIITTITPGAFTTLVSLSTINLLSNPFNCNCHLAWLGKWLKRKR  
 IVSGNPRCQKPFFLKEIPIQDVAIQDFTCDGNEESSCQLSPRCPEQCTCMETVVRCSNKGLR  
 ALPRGMPKDVTELYLEGNHLTAVPRELSALRHLTLIDLSNNSISMLTNYTFSNMHLSTLIL  
 SYNRLRCIPVHAFNGLRSLRVLTTLHGNDISSVPEGSFNDLTSLSHLALGTNPLHCDCLRWL  
 SEWVKAGYKEPGIARCSSPEPMADRLLLTTPTHRFQCKGPVDINIVAKCNACLSSPCKNNGT  
 CTQDPVELYRCACPYSYKGKDVTPINTCIQNPCQHGGTCHLSDSHKDGFS CSCPLGFEGQR  
 CEINPDDCEDNDCENNATCVDGINNYVCICPPNYTGELCDEVIDHCVPELNLQHEAKCIPL  
 DKGFSCECVPGYSGKLCETDNDDCVAHKCRHGAQCVDITINGYTCTCPQGFSGPFCEHPPPMV  
 LLQTSPCDQYECQNGAQCIVVQQEPTCRCPPGFAGPRCEKLITVNFVGKDSYVELASAKVRP  
 QANISLQVATDKDNGILLYKGDNDPLALELYQGHVRLVYDSLSSPPTTVYSVETVNDGQFHS  
 VELVTLNQTLNLVVDKGTPKSLGKLQKQPAVGINSPLYLGGIPTSTGLSALRQGTDRPLGGF  
 HGCIHEVRINNELQDFKALPPQSLGVSPGCKSCTVCKHGLCRSVEKDSVCECRPGWTGPLC  
 DQEARDPCLGHRCHHGKCVATGTSYMCKCAEGYGGDLCDNKND SANACSAFKCHHGQCHISD  
 QGEPYCLCQPGFSGEHCQQENPCLGQVVREVIRRQKGYASCATASKVPIMECRGGCGPQCCQ  
 PTRSKRRKYVFQCTDGSSFVEEVERHLECGCLACS

**Signal peptide:**

amino acids 1-27

**FIGURE 113**

GGATGCAGGACGCTCCCCTGAGCTGCCTGTCACCGACTAGGTGGAGCAGTGTTTCTTCCGCA  
GACTCAACTGAGAAGTCAGCCTCTGGGGCAGGCACCAGGAATCTGCCTTTTCAGTTCTGTCT  
CCGGCAGGCTTTGAGG**ATGA**AGGCTGCGGGCATTCTGACCCTCATTGGCTGCCTGGTCACAG  
GCGCCGAGTCCAAAATCTACACTCGTTGCAAACCTGGCAAAAATATTCTCGAGGGCTGGCCTG  
GACAATTACTGGGGCTTCAGCCTTGAAACTGGATCTGCATGGCATATTATGAGAGCGGCTA  
CAACACCACAGCCCCGACGGTCCTGGATGACGGCAGCATCGACTATGGCATCTTCCAGATCA  
ACAGCTTCGCGTGGTGCAGACGCGGAAAGCTGAAGGAGAACAACCACTGCCATGTCGCCTGC  
TCAGCCTTGATCACTGATGACCTCACAGATGCAATTATCTGTGCCAGGAAAATTGTTAAAGA  
GACACAAGGAATGAACTATTGGCAAGGCTGGAAGAAACATTGTGAGGGCAGAGACCTGTCCG  
AGTGGA AAAAAGGCTGTGAGGTTTCC**TAA**ACTGGAACCTGGACCCAGGATGCTTTGCAGCAAC  
GCCCTAGGATTTGCAGTGAATGTCCAAATGCCTGTGTCATCTTGTCCCGTTTCCTCCCAATA  
TTCCTTCTCAAACCTTGAGAGGGAAAATTAAGCTATACTTTTAAGAAAATAAATATTTCCAT  
TTAAATGTC

## **FIGURE 114**

MKAAGILTLIGCLVTGAESKIYTRCKLAKIFSRAGLDNYWGFSLGNWICMAYYESGYNTTAP  
TVLDDGSIDYGIFQINSAWCRGKCLKENNHCHVACSALITDDLTDALICARKIVKETQGMN  
YWQGWKKHCEGRDLSEWKKGCEVS

**Signal peptide:**

amino acids 1-19

115/249

## FIGURE 115

[illegible]



**FIGURE 116**

MPLLKLVHGSPLVFGEKFKLFTLVSACIPVFERLARRRKKILFYCHFDPDLLLTKRDSFLKRLY  
RAPIDWIEEYTTGMADCILVNSQFTAAVFKETFKSLSHIDPDVLYPSLNVTSFDSVVPEKLD  
DLVPKGKKFLLLSINRYERKKNLTLALEALVQLRGRLTSQDWERVHLIVAGGYDERVLENVE  
HYQELKKMVQQSDLGQYVTFILRSFSDKQKISLLHSCTCVLYTPSNEHFGIVPLEAMYMQCPV  
IAVNSSGGPLESIDHSVTGFLCEPDPVHFSEAIEKFIREPSLKATMGLAGRARVKEKFSPEAF  
TEQLYRYVTKLLV

**Signal peptide:**

amino acids 1-15

117/249

**FIGURE 117**

GACTACGCCGATCCGAGACGTGGCTCCCTGGGCGGCAGAACC**ATG**TTGGACTTCGCGATCTT  
 CGCCGTTACCTTCTTGCTGGCGTTGGTGGGAGCCGTGCTCTACCTCTATCCGGCTTCCAGAC  
 AAGCTGCAGGAATTCCAGGGATTACTCCAACCTGAAGAAAAAGATGGTAATCTTCCAGATATT  
 GTGAATAGTGGAAGTTTGCATGAGTTCCTGGTTAATTTGCATGAGAGATATGGGCCTGTGGT  
 CTCCTTCTGGTTTGGCAGGCGCCTCGTGGTTAGTTTGGGCACTGTTGATGTACTGAAGCAGC  
 ATATCAATCCCAATAAGACATCGGACCCTTTTGAAACCATGCTGAAGTCATTATTAAGGTAT  
 CAATCTGGTGGTGGCAGTGTGAGTGAAAACCATGAGGAAAAAATTGTATGAAAATGGTGT  
 GACTGATTCTCTGAAGAGTAACCTTTGCCCTCCTCCTAAAGCTTTCAGAAGAATTATTAGATA  
 AATGGCTCTCCTACCCAGAGACCCAGCACGTGCCCCCTCAGCCAGCATATGCTTGGTTTTGCT  
 ATGAAGTCTGTTACACAGATGGTAATGGGTAGTACATTTGAAGATGATCAGGAAGTCATTCTG  
 CTTCCAGAAGAATCATGGCACAGTTTGGTCTGAGATTGGAAAAGGCTTTCTAGATGGGTCTAC  
 TTGATAAAAACATGACTCGGAAAAACAATATGAAGATGCCCTCATGCAACTGGAGTCTGTT  
 TTAAGGAACATCATAAAAGAACGAAAAGGAAGGAACCTTCAGTCAACATATTTTTCATTGACTC  
 CTTAGTACAAGGGAACCTTAATGACCAACAGATCCTAGAAGACAGTATGATATTTTCTCTGG  
 CCAGTTGCATAATAACTGCAAAATTGTGTACCTGGGCAATCTGTTTTTTAACCACCTCTGAA  
 GAAGTTCAAAAAAATTATATGAAGAGATAAACCAAGTTTTTGGAAATGGTCCTGTACTCC  
 AGAGAAAATTGAGCAGCTCAGATATTGTCAGCATGTGCTTTGTGAAACTGTTTCGAACTGCCA  
 AACTGACTCCAGTTTCTGCCCAGCTTCAAGATATTGAAGGAAAAATTGACCGATTTATTATT  
 CCTAGAGAGACCCTCGTCCTTTATGCCCTTGGTGTGGTACTTCAGGATCCTAATACTTGGCC  
 ATCTCCACACAAGTTTGATCCAGATCGGTTTGTATGATGAATTAGTAATGAAAACTTTTTCCT  
 CACTTGGATTCTCAGGCACACAGGAGTGTCCAGAGTTGAGGTTTGCATATATGGTGACCACA  
 GTACTTCTTAGTGTATTGGTGAAGAGACTGCACCTACTTTCTGTGGAGGGACAGGTTATTGA  
 AACAAAGTATGAACTGGTAACATCATCAAGGGAAGAAGCTTGGATCACTGTCTCAAAGAGAT  
 AT**TAA**AATTTTATACATTTAAATCATTTGTTAAATTGATTGAGGAAAAACAACCATTTAAAAA  
 AAATCTATGTTGAATCCTTTTATAAACAGTATCACTTTGTAATATAAACACCTATTTGTAC  
 TTAA

**FIGURE 118**

MLDFAIFAVTFLLALVGAVLYLYPASRQAAGIPGITPTEEKDGNLPDIVNSGSLHEFLVNLH  
ERYGPVVSFWFGRRLVVS LGTVDLKQHINPNKTSDFETMLKSLLRYQSGGGSVSENHMRK  
KLYENGVTDSLKSNFALLKLSEELLDKWLSYPETQHVPLSQHMLGFAMKSVTQMVMGSTFE  
DDQEVIRFQKNHGTWVSEIGKGFLDGS LDKNMTRKKQYEDALMQLESVLRNIIKERKGRNFS  
QHIFIDSLVQGNLNDQQILED SMIFSLASCIITAKLCTWAICFLTTS EEVQKKLYEEINQVF  
GNGPVTPEKIEQLRYCQHVL CETVRTAKLTPVSAQLQDIEGKIDRFIIPRETLVLYALGVVL  
QDPNTWPSPHKFDPDRFDDELVMKTFSSLGFSGTQECPELRFAYMVT TVLLSVLVKRLHLLS  
VEGQVIETKYELVTSSREEAWITVSKRY

**Signal peptide:**

amino acids 1-18

**Transmembrane domain:**

amino acids 271-290

**FIGURE 119**

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGA  
GACCGCCGCCCTTGTCCCCGAGGGCC**ATG**GGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTC  
CTGACGCTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACAT  
ACAGGCCTGCCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTTCAGCTGG  
TGGCCGCGCTCTCTGTACCCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGA  
GTCTCCATGTTCAACAGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGT  
GGCCCTGTCCTTCTTCATATTCGAGCGTTGGGAGTGCACCTACGTATTGGTACATTTTTGTCT  
TCTGCAGTGCCCTTCCAGCTGTCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAA  
AAGAAACCCTTCT**TGA**ATTACCTTCATGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGG  
CCGCTTCGTATTCCTGGAAGAAGGAAGGCATAGGCTTCGGTTTTCCCTCGGAAACTGCTTC  
TGCTGGAGGATATGTGTTGGAATAATTACGTCTTGAGTCTGGGATTATCCGCATTGTATTTA  
GTGCTTTGTAATAAAATATGTTTTGTAGTAACATTAAGACTTATATACAGTTTTAGGGGACA  
ATTAAAAAAAAAAAAA

## **FIGURE 120**

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLG  
LFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVT  
EMALFVTVFGLKKKPF

### **Transmembrane domain:**

amino acids 12-28 (type II), 51-66, 107-124

121/249

**FIGURE 121**

TCCCGGACCCTGCCGCCCTGCCACT**ATG**TCCCGCCGCTCTATGCTGCTTGCCTGGGCTCTCC  
CCAGCCTCCTTCGACTCGGAGCGGCTCAGGAGACAGAAGACCCGGCCTGCTGCAGCCCCATA  
GTGCCCCGGAACGAGTGGAAGGCCCTGGCATCAGAGTGCGCCCAGCACCTGAGCCTGCCCTT  
ACGCTATGTGGTGGTATCGCACACGGCGGGCAGCAGCTGCAACACCCCCGCCTCGTGCCAGC  
AGCAGGCCCGGAATGTGCAGCACTACCACATGAAGACACTGGGCTGGTGCAGCTGGGCTAC  
AACTTCCTGATTGGAGAAGACGGGCTCGTATACGAGGGCCGTGGCTGGAACCTCACGGGTGC  
CCACTCAGGTCACCTTATGGAACCCCATGTCCATTGGCATCAGCTTCATGGGCAACTACATGG  
ATCGGGTGCCACACCCCAGGCCATCCGGGCAGCCCAGGGTCTACTGGCCTGCGGTGTGGCT  
CAGGGAGCCCTGAGGTCCAACCTATGTGCTCAAAGGACACCGGGATGTGCAGCGTACACTCTC  
TCCAGGCAACCAGCTCTACCACCTCATCCAGAATTGGCCACACTACCGCTCCCCC**TG**AGGCC  
CTGCTGATCCGCACCCCATTCCTCCCCTCCCATGGCCAAAAACCCCACTGTCTCCTTCTCCA  
ATAAAGATGTAGCTC

**FIGURE 122**

MSRRSMLLAWALPSLLRLGAAQETEDPACCSPIVPRNEWKALASECAQHLSLPLRYVVVSHT  
AGSSCNTPASCQQQARNVQHYHMKTLGWCDVGYNFLIGEDGLVYEGRGWNFTGAHSGHLWNP  
MSIGISFMGNYMDRVPTPQAIRAAQGLLACGVAQGALRSNYVLKGHRDVQRTLSPGNQLYHL  
IQNWPHYRSP

**Signal peptide:**

amino acids 1-20

**FIGURE 123**

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAG**ATG**CAACT  
GACTCGCTGCTGCTTCGTGTTCTTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGG  
ATGATGGTCCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGGCCAGCCCCGGCCC  
CGGGTGCCTCGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCT  
CCTAGGGCTGCTGGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCC  
CGAACCACAGCCCCCACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTAC  
TCCAACATCAAGACGGTGGCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAA  
TGGGACCTTCAGCGTCCACTTCCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCC  
TCGTGCCCCCAGTAAAGCTGTAGAGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAG  
GCCTCCAAAATCTTCAACTGCCGGATGGAGTGGGAGAAGGTAGAACGGGGCCGCCGGACCTC  
GCTTTGCACCCACGACCCAGCCAAGATCTGCTCCCGAGACCACGCTCAGAGCTCAGCCACCT  
GGAGCTGCTCCCAGCCCTTCAAAGTCGTCTGTGTCTACATCGCCTTCTACAGCACGGACTAT  
CGGCTGGTCCAGAAGGTGTGCCCAGATTACAACTACCATAGTGATACCCCCTACTACCCATC  
TGGG**TGA**CCCGGGGCAGGCCACAGAGGCCAGGCCAGGGCTGGAAGGACAGGCCTGCCCATGC  
AGGAGACCATCTGGACACCGGGCAGGGAAGGGGTTGGGCCTCAGGCAGGGAGGGGGGTGGAG  
ACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAGAGAAAGGGTCCCAAGTG  
CTGGTCCCAACCTGAAGCTGTGGAGTGAAGTAGATCACAGGAGCACTGGAGGAGGAGTGGGCT  
CTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGGTCCCCGAGGCC  
TGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGCCCTTGTTT  
CTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACAGCCTG  
TCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGGGC  
CAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCCTTGTCGTGTGCTGAGCATG  
GCATGAGGCTGAAGTGGCAACCCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTC  
CAGCCAGGCCACCCCTTTCCAAAATTCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATT  
GCTGATGGCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCCTCCACACTAAGGCC  
ACAGCCCATCCGCGTGCTGTGTGTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAG  
CATCCATGTCCCGGAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTCAGACCGGGGTTCTCC  
CGGATCTGGATGGCGCCGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGGCCGAGA  
GCATGTGCTGGATCTGTTCTGTGTGTCTGTCTGTGGGTGGGGGGAGGGGAGGGAAGTCTTGT  
GAAACCGCTGATTGCTGACTTTTGTGTGAAGAATCGTGTTCTTGGAGCAGGAAATAAGCTT  
GCCCCGGGGCA



**FIGURE 124**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66521

><subunit 1 of 1, 252 aa, 1 stop

><MW: 28127, pI: 8.91, NX(S/T): 5

MQLTRCCFVFLVQGS<sup>1</sup>LYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMAN  
STLLGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVD  
HGNGTFSVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGR  
RTSLCTHDPAKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVC PDYNYHSDTPY  
YPSG

**Important features of the protein:**

**Signal peptide:**

amino acids 1-14

**N-glycosylation sites.**

amino acids 62-65, 127-130, 137-140, 143-146

**2-oxo acid dehydrogenases acyltransferase**

amino acids 61-71

**FIGURE 125**

GTGAATGTGAGGGTTTGATGACTTTCAGATGTCTAGGAACCAGAGTGGGTGCAGGGGCCCCA  
 GGCAGGGCTGATTCTTGGGCGGAGGAGAGTAGGGTAAAGGGTTCTGCATGAGCTCCTTAAAG  
 GACAAAGGTAACAGAGCCAGCGAGAGAGCTCGAGGGGAGACTTTGACTTCAAGCCACAGAAT  
 TGGTGGAAAGTGTGCGCGCCGCCGCCGCTCGCTCCTGCAGCGCTGTGACCTAGCCGCTAG  
 CATCTTCCCGAGCACCGGGATCCCGGGGTAGGAGGCGACGCGGGCGAGCACCAGCGCCAGCC  
 GGCTGCGGCTGCCCACACGGCTCACC**ATG**GGGCTCCGGGCGCCGGGCGCTGTCCGCGGTGCCG  
 GCCGTGCTGCTGGTCCTCACGCTGCCGGGGCTGCCCGTCTGGGCACAGAACGACACGGAGCC  
 CATCGTGCTGGAGGGCAAGTGTCTGGTGGTGTGCGACTCGAACCCGGCCACGGACTCCAAGG  
 GCTCCTCTTCTCCCCGCTGGGGATATCGGTCCGGGCGGCCAACTCCAAGGTCGCCTTCTCG  
 GCGGTGCGGAGCACCACCACGAGCCATCCGAGATGAGCAACAAGACGCGCATCATTTACTT  
 CGATCAGATCCTGGTGAATGTGGGTAAATTTTTTACATTGGAGTCTGTCTTTGTAGCACCAA  
 GAAAAGGAATTTACAGTTTCAGTTTTTACGTGATTAAAGTCTACCAGAGCCAACTATCCAG  
 GTTAACTTGATGTTAAATGGAAAACCAGTAATATCTGCCTTTGCGGGGGACAAAGATGTTAC  
 TCGTGAAGCTGCCACGAATGGTGTCTCTACCTAGATAAAAGAGGATAAGGTTTACCTAA  
 AACTGGAGAAAGGTAATTTGGTTGGAGGCTGGCAGTATTCACGTTTTCTGGCTTTCTGGTG  
 TTCCCCCTA**TAG**GATTCAATTTCTCCATGATGTTTCATCCAGGTGAGGGATGACCCACTCCTG  
 AGTTATTGGAAGATCATTTTTTTCATCATTGGATTGATGTCTTTTATTGGTTTCTCATGGGTG  
 GATATGGATTCTAAGGATTCTAGCCTGTCTGAACCAATACAAAATTTACAGATTATTTGTG  
 TGTGTCTGTTTCAGTATATTTGGATTGGGACTCTAAGCAGATAATACCTATGCTTAAATGTA  
 ACAGTCAAAAGCTGTCTGCAAGACTTATTCTGAATTTCAATTCCTGGGATTACTGAATTAGT  
 TACAGATGTGGAATTTTATTTGTTTAGTTTTTAAAAGACTGGCAACCAGGTCTAAGGATTAGA  
 AAACCTCTAAAGTTCTGACTTCAATCAACGGTTAGTGTGATACTGCCAAAGAACTGTATACTG  
 TGTTAATATATTGATTATATTTGTTTTTATTCCTTTGGAATTAGTTTGTTTGGTTCTTGTA  
 AAAACTTGGAATTTTTTTTTTTCAGTAACTGGTATTATGTTTTCTCTTAAATAAGGTAATGAA  
 TGGCTTGCCACAAATTTACCTTGACTACGATATCATCGACATGACTTCTCTCAAAAAAAAAA  
 GAATGCTTCATAGTTGTATTTTAATTGTATATGTGAAAGAGTCATATTTTCCAAGTTATATT  
 TTCTAAGAAGAAGATAGATCATAAATCTGACAAGGAAAAAGTTGCTTACCCAAAATCTAAG  
 TGCTCAATCCCTGAGCCTCAGCAAAACAGCTCCCCCTCCGAGGGAAATCTTATACTTTATTGC  
 TCAACTTTAATTAAAATGATTGATAATAACCACTTTATTAAAAACCTAAGGTTTTTTTTTTT  
 TCCGTAGACATGACCACTTTATTAACTGGTGGTGGGATGCTGTTGTTTCTAATTATACCTAT  
 TTTTCAAGGCTTCTGTTGTATTTGAAGTATCATCTGGTTTTGCCTTAACTCTTTAAATTGTA  
 TATATTTATCTGTTTAGCTAATATTAAATTCAAATATCCCATATCTAAATTTAGTGCAATAT  
 CTTGTCTTTTGTATAGGTCATATGAATTCATAAAATTATTTATGTCTGTTATAGAATAAAGA  
 TTAATATATGTTAAAAAAA

**FIGURE 126**

MSGGRRALSAVPAVLLVLTLPGLPVWAQNDTEPIVLEGKCLVVCDSNPATDSKGSSSSPLGI  
SVRAANSKVAFSAVRSTNHEPSEMSNKTRIIYFDQILVNVGNFFTLESV FVAPRKGIYSFSF  
HVIKVYQSQTIQVNLMLNGKPVISAFAGDKDVTREAAATNGVLLYLDKEDKVYLKLEKGNLVG  
GWQYSTFSGFLVFPL

**Signal peptide:**

amino acids 1-27

127/249

**FIGURE 127**

CGGTGGCC**ATG**ACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTC  
GCCCTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGC  
TTTCTTCTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTA  
TTGACAACAAAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTC  
TATATCCAAGAAATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTT  
GAAGAGTATAAAACCCAGGTGAGACAGCACCTCTATGCGACTGCTGGCCTATGTTTCTGGCT  
TGGGCTTTGGAATCATGAGTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGG  
CCAGGCACAGTGGGCATTTCATGGAGATTCTCCTCAATTCTTCCTTTATTCAGCTTTCATGAC  
GCTGGTCATTATCTTGCTGCATGTATTCTGGGGCATTGTATTTTTTTGATGGCTGTGAGAAGA  
AAAAGTGGGGCATCCTCCTTATCGTTCTCCTGACCCACCTGCTGGTGTGAGCCAGACCTTC  
ATAAGTTCTTATTATGGAATAAACCTGGCGTCAGCATTTATAATCCTGGTGCTCATGGGCAC  
CTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCGAAGCCTGAAACTCTGCCTGCTCTGCCAAG  
ACAAGAACTTTCTTCTTTACAACCAGCGCTCCAGAT**TAA**CCTCAGGGAACCAGCACTTCCCAA  
ACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCTTTTTTCTGAAAATCCCTTTTTTCTG  
GTGGAATTGAGAAAGAAATAAACTATGCAGATA

**FIGURE 128**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66658

><subunit 1 of 1, 257 aa, 1 stop

><MW: 28472, pI: 9.33, NX(S/T): 0

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDN  
 KDGPQTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLG  
 GIMSGVFSFVNLTSLGPGTVGIHGDSPOFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKW  
 GILLIVLLTHLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSCLKCLLCQDKN  
 FLLYNQRSR

**Important features of the protein:**

**Signal peptide:**

amino acids 1-19

**Transmembrane domains:**

amino acids 32-51, 119-138, 152-169, 216-235

**Glycosaminoglycan attachment site.**

amino acids 120-123

**Sodium:neurotransmitter symporter family protein**

amino acids 31-65



**FIGURE 130**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66659
><subunit 1 of 1, 832 aa, 1 stop
><MW: 94454, pI: 6.94, NX(S/T): 12
MFALGLPFLVLLVASVESHLGVLGPKNVSQKDAEFERTYVDEVNSELVNIYTFNHTVTRNRT
EGVRVSVNVLNKQKGAPLLFVVRQKEAVVSFQVPLILRGMFQRKYLYQKVERTLCQPPTKNE
SEIQFFYVDVSTLSPVNTTYQLRVSRMDDFVLRTGEQFSFNTTAAQPQYFKYEFPEGVDSVI
VKVTSNKAFPCSVISIQDVLCPVYDLNNAVFIGMYQTMTKKAAITVQRKDFPSNSFYVVVV
VKTEDQACGGSPLPFYPFAEDEPVDQGHRQKTL SVLV SQAVTSEAYVSGMLFCLGIFLSFYLL
TVLLACWENWRQKKKTLLVAIDRACPESGHPRLADSFPGSSPYEGYNYGSFENVSGSTDGL
VDSAGTGDLSYGYQGRSFEPVGTRPRVDSMSSVEEDDYDTLTDIDSDKNVIRTKQYLYVADL
ARKDKRVLRRKKYQIYFWNIATIAVFYALPVVQLVITYQTVVNVVTGNQDICYYNFLCAHPLGN
LSAFNNILSNLGYILLGLLFLLIILQREINHNRRALLRNDLCALECGIPKHFGFLFYAMGTALM
MEGLLSACYHVCPNYTNFQFDTSFMYMIAGLCMLKLYQKRHPDINASAYSAYACLAIVIFFS
VLGVVFGKGNTAFWIVFSIIHIIATLLLSTQLYYMGRWKLDSGIFRRILHVLYTDCIRQCSG
PLYVDRMVLLVMGNVINWSLAAYGLIMRPNDFASYLLAIGICNLLLYFAFYIIMKLRSGERI
KLIPLLCIVCTSVVWGFALFFFFQGLSTWQKTPAESREHNRDCILLDFFDDHDIWHFLSSIA
MFGSFLVLLTLDDDLDTVQRDKIYVF
```

**Important features of the protein:****Signal peptide:**

amino acids 1-18

**Transmembrane domains:**

amino acids 292-317, 451-470, 501-520, 607-627, 751-770

**Leucine zipper pattern.**

amino acids 497-518

**N-glycosylation sites.**

amino acids 27-30, 54-57, 60-63, 123-126, 141-144, 165-168, 364-367, 476-479, 496-499, 572-575, 603-606, 699-702

**FIGURE 131**

GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGC  
 TCTCTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCCTG  
 ACC**ATG**GTCCTTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCCAGGCTCTCCCCAAGGC  
 CCAGCCTGCAGAGCTGTCTGTGGAAGTTCCAGAAAACCTATGGTGGAATTTCCCTTTATACC  
 TGACCAAGTTGCCGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTCAGGGGACTCA  
 GGCAAGGCAACTGAGGGGCCATTTGCTATGGATCCAGATTCTGGCTTCCTGCTGGTGACCAG  
 GGCCCTGGACCGAGAGGAGCAGGCAGAGTACCAGCTACAGGTACCCCTGGAGATGCAGGATG  
 GACATGTCTTGTGGGGTCCACAGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTG  
 CCCCATTCTCTCAAGCCATCTACAGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCC  
 CTTCTCTTCTTGAGGCTTCAGACCGGGATGAGCCAGGCACAGCCAACTCGGATCTTCGAT  
 TCCACATCCTGAGCCAGGCTCCAGCCCAGCCTTCCCCAGACATGTTCCAGCTGGAGCCTCGG  
 CTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGCACCAGCCTTGACCACGCCCTGGAGAGGAC  
 CTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGACCAGGCCTCAGGCCACCAGGCCACTG  
 CCACCGTGGAAGTCTCCATCATAGAGAGCACCTGGGTGTCCCTAGAGCCTATCCACCTGGCA  
 GAGAATCTCAAAGTCTATACCCGCACCACATGGCCCAGGTACACTGGAGTGGGGGTGATGT  
 GCACTATCACCTGGAGAGCCATCCCCCGGGACCCCTTTGAAGTGAATGCAGAGGGAAACCTCT  
 ACGTGACCAGAGAGCTGGACAGAGAAGCCCAGGCTGAGTACCTGCTCCAGGTGCGGGCTCAG  
 AATTCCTATGGCGAGGACTATGCGGGCCCTCTGGAGCTGCACGTGCTGGTGATGGATGAGAA  
 TGACAACGTGCCTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCTCAGTCCAC  
 CAGGTACTGAAGTGAAGTACTGTGAGCAGAGGATGCAGATGCCCCCGGCTCCCCCAATTCC  
 CACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTCCA  
 GGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACA  
 TCCTGCTTCTGGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTGT  
 GAAGTGAAGTCGCAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTTCCCAGAT  
 TGGGCCTATAAGCCTCCCTGAGGATGTGGAGCCCGGGACTCTGGTGGCCATGCTAACAGCCA  
 TTGATGCTGACCTCGAGCCCGCCTTCCGCCTCATGGATTTTGCCATTGAGAGGGGAGACACA  
 GAAGGGACTTTTGGCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAA  
 GAACCTCAGTTATGAGGCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGA  
 AGCTGGTGGGGCCAGGCCCAGGCCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGGAGAGA  
 GTGATGCCACCCCCAAGTTGGACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCCC  
 AGCCGGCTCTTTCTGCTGACCATCCAGCCCTCCGACCCCATCAGCCGAACCTCAGGTTCT  
 CCCTAGTCAATGACTCAGAGGGCTGGCTCTGCATTGAGAAATTCTCCGGGGAGGTGCACACC  
 GCCAGTCCCTGCAGGGCGCCCAGCCTGGGGACACCTACACGGTGCTTGTGGAGGGCCAGGA  
 TACAGCCCTGACTCTTGCCCCTGTGCCCTCCCAATACCTCTGCACACCCCGCCAAGACCATG  
 GCTTGATCGTGAGTGGACCCAGCAAGGACCCCGATCTGGCCAGTGGGCACGGTCCCTACAGC  
 TTCACCCTTGGTCCCAACCCACGGTGCAACGGGATTGGCGCCTCCAGACTCTCAATGGTTC  
 CCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGAGCCACGTGAACACATAATCCCCGTGG  
 TGGTCAGCCACAATGCCAGATGTGGCAGCTCCTGGTTCGAGTGATCGTGTGCTGCAAC  
 GTGGAGGGGAGTGCATGCGCAAGGTGGGCCGCATGAAGGGCATGCCACGAAGCTGTCCGC  
 AGTGGGCATCCTTGTAGGCACCCTGGTAGCAATAGGAATCTTCCTCATCCTCATTTTACCC  
 ACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCCTGAAGGCG  
 ACTGTCT**TGA**ATGGCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAGTCCC  
 CTGGGAGAGAGCCAGCACCCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCCCTCCAT  
 CTGCCCTGGGGTGGAGGCACCATCACCATCACCAGGCATGTCTGCAGAGCCTGGACACCAAC  
 TTTATGGACTGCCCATGGGAGTGCTCCAAATGTGAGGGTGTTTGCCCAATAATAAAGCCCCA  
 GAGAACTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG



**FIGURE 132**

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSG  
KATEGPFAMDPDSGFLLVTRALDREEQAQYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVP  
HFSQAIYRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSDFMFQLEPRL  
GALALSPKGSTSLDHALERTYQLLVQVKMDQASGHQATATVEVSIIESTWVSLEPIHLAE  
NLKVLYPHHMAQVHWSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQN  
SHGEDYAAPLELHVLVMDENDNVPICPPRDPTVSIPELSPPGTEVTRLAEDADAPGSPNSH  
VVYQLLSPEPEDGVEGRAFQVDPTSGSVTLGVLPLRAGQNILLVLAMDLAGAEGGFSSTCE  
VEVAVTDINDHAPEFITSQIGPISLPEDVEPGTLVAMLTAIDADLEPAFRLMDFAIERGDTE  
GTFGLDWEPSGHVRLRLCKNLSYEAAPSHEVVVVVQSVAKLVGPGPGPGATATVTVLVERV  
MPPPKLDQESYEASVPISAPAGSFLLTIQPSDPISRTLRFSLVNDSEGWLCIEKFSGEVHTA  
QSLQGAQPGDTYTVLVEAQDTALTAPVPSQYLCTPRQDHGLIVSGPSKDPDLASGHGPYSF  
TLGPNPTVQRDWRLQTLNGSHAYLTALHWWVEPREHIIPVVVSHNAQMWQLLVRVIVCRCNV  
EGQCMRKVGRMKGMPTKLSAVGILVGTLLVAIGIFLILIFTHWTMSRKKDPDQPADSVPLKATV

**Signal peptide:**

amino acids 1-18

**Transmembrane domain:**

amino acids 762-784

**FIGURE 133**

CCGGGGAC**ATG**AGGTGGATACTGTTCAATTGGGGCCCTTATTGGGTCCAGCATCTGTGGCCAA  
 GAAAAATTTTTTGGGGACCAAGTTTTTGAGGATTAATGTCAGAAATGGAGACGAGATCAGCAA  
 ATTGAGTCAACTAGTGAATTCAAACAACCTGAAGCTCAATTTCTGGAAATCTCCCTCCTCCT  
 TCAATCGGCCTGTGGATGTCCTGGTCCCATCTGTCACTCTGCAGGCATTTAAATCCTTCCTG  
 AGATCCCAGGGCTTAGAGTACGCAGTGACAATTGAGGACCTGCAGGCCCTTTTAGACAATGA  
 AGATGATGAAATGCAACACAATGAAGGGCAAGAACGGAGCAGTAATAACTTCAACTACGGGG  
 CTTACCATTCCCTGGAAGCTATTTACCACGAGATGGACAACATTGCCGCAGACTTTTCCTGAC  
 CTGGCGAGGAGGGTGAAGATTGGACATTCGTTTGAACCGGCCGATGTATGTACTGAAGTT  
 CAGCACTGGGAAAGGCGTGAGGCGGCCGGCGTTTGGCTGAATGCAGGCATCCATTCCCAG  
 AGTGATCTCCCAGGCCACTGCAATCTGGACGGCAAGGAAGATTGTATCTGATTACCAGAGG  
 GATCCAGCTATCACCTCCATCTTGGAGAAAAATGGATATTTTCTTGTGCTGTGGCCAATCC  
 TGATGGATATGTGTATACTCAAACCTCAAACCGATTATGGAGGAAGACGCGGTCCCAGAAATC  
 CTGGAAGCTCCTGCATTGGTGCTGACCCAAATAGAACTGGAACGCTAGTTTTGTCAGGAAAG  
 GGAGCCAGCGACAACCCTTGCTCCGAAGTGATCCATGGACCCACGCCAATTCGGAAGTGGA  
 GGTGAAATCAGTGGTAGATTTTCATCCAAAAACATGGGAATTTCAAGGGCTTCATCGACCTGC  
 ACAGCTACTCGCAGCTGCTGATGTATCCATATGGGTACTCAGTCAAAAAGGCCCCAGATGCC  
 GAGGAACTCGACAAGGTGGCGAGGCTTGCGGCCAAAGCTCTGGCTTCTGTGTCGGGCACTGA  
 GTACCAAGTGGGTCCCACCTGCACCACTGTCTATCCAGCTAGCGGGAGCAGCATCGACTGGG  
 CGTATGACAACGGCATCAAATTTGCATTCACATTTGAGTTGAGAGATAACGGGACCTATGGC  
 TTCCTCCTGCCAGCTAACCAGATCATCCCCACTGCAGAGGAGACGTGGCTGGGGCTGAAGAC  
 CATCATGGAGCATGTGCGGGACAACCTCTAC**TAG**GCGATGGCTCTGCTCTGTCTACATTTAT  
 TTGTACCCACACGTGCACGCACTGAGGCCATTGTTAAAGGAGCTCTTTCCTACCTGTGTGAG  
 TCAGAGCCCTCTGGGTTTGTGGAGCACACAGGCCTGCCCTCTCCAGCCAGCTCCCTGGAGT  
 CGTGTGTCCTGGCGGTGTCCCTGCAAGAACTGGTTCTGCCAGCCTGCTCAATTTTGGTFCCTG  
 CTGTTTTTGTATGAGCCTTTTGTCTGTTTCTCCTTCCACCCTGCTGGCTGGGCGGCTGCACTC  
 AGCATACCCCTTTCCTGGGTGGCATGTCTCTCTCTACCTCATTTTTTAGAACCAGAAACATC  
 TGAGATGATTCTCTACCTCATCCACATCTAGCCAAGCCAGTGACCTTGCTCTGGTGGCACT  
 GTGGGAGACACCACTTGTCTTTAGGTGGGTCTCAAAGATGATGTAGAATTTCTTTAATTTT  
 TCGCAGTCTTCTTGGAAAATATTTTCTTTTGGAGCAGCAAATCTTGTAGGGATATCAGTGAAG  
 GTCTCTCCCTCCCTCCTCTCCTGTTTTTTTTTTTTTGGAGACAGAGTTTTGCTCTTGTGTC  
 CAGGCTGGAGTGTGATGGCTCGATCTTGGCTCACCAACCTCTGCCTCCTGGGTTCAGGA  
 ATTCTCCTGCCTCAGCCTCTTGAGTAGCTTGGTTTATAGGCGCATGCCACCATGCCTGGCTA  
 ATTTTGTGTTTTTAGTAGAGACAGGGTTTCTCCATGTTGGTCAGGCTGGTCTCAAACCTCCCA  
 ACCTCAGGTGATCTGCCCTCCTTGGCCTCCCAGAGTGCTGGGATTACAGGTGTGAGCCACTG  
 TGCCGGGCCCCGTCCCCTCCTTTTTTAGGCCTGAATACAAAGTAGAAGATCACTTTCCTTCAC  
 TGTGCTGAGAATTTCTAGATACTACAGTTCTTACTCCTCTCTTCCCTTTGTTATTACAGTGTG  
 ACCAGGATGGCGGGAGGGGATCTGTGTCACTGTAGGTACTGTGCCAGGAAGGCTGGGTGAA  
 GTGACCATCTAAATTGCAGGATGGTGAAATTATCCCCATCTGTCCTAATGGGCTTACCTCCT  
 CTTTGCCTTTTGAACCTCACTTCAAAGATCTAGGCCTCATCTTACAGGTCTAAATCACTCAT  
 CTGGCCTGGATAATCTCACTGCCCTGGCACATTCCCATTTGTGCTGTGGTGTATCCTGTGTT  
 TCCTTGTCTGGTTTG  
 TCTGTCTATTTTGTATCCTGGACCACAAGTTCCTAAGTAGAGCAAGAATTCATCAACCAGCT  
 GCCTCTTGTTCATTTACCTCAGCACGTACCATCTGTCTTTTGTGTTGTGTTGTTGTTGTTT  
 TTGTTTTTTTGTCTTTTACCAAACATGTCTGTAAATCTTAACCTCCTGCCTAGGATTTGTACA  
 GCATCTGGTGTGTGCTTATAAGCCAATAAATATTCAATGTGAAAAAAAAAAAAAAAAAAAA

**FIGURE 134**

MRWILFIGALIGSSICGQEKFFGDQVLRINVRNGDEISKLSQLVNSNNLKLNFWKSPSSFNR  
PVDVLVPSVSLQAFKSFLRSQGLEAYVTIEDLQALLDNEDDEMQHNEGQERSNNFNYGAYH  
SLEAIYHEMDNIAADFPDLARRVKIGHSFENRPMYVLKFSTGKGVRRPAVWLNAGIHSREWI  
SQATAIWTARKIVSDYQRDPAITSILEKMDIFLLPVANPDGYVYTQTQNRLLWRKTRSRNPGS  
SCIGADPNRNWNASFAGKGASDNPCSEVYHGPHANSEVEVKSVVDFIQKHGNFKGFIDLHSY  
SQLLMYPYGYSVKKAPDAEELDKVARLAAKALASVSGTEYQVGPTCTTVYPASGSSIDWAYD  
NGIKFAFTFELRDTGTYGFLLPANQIIPAEETWLGLKTIMEHVRDONLY

**Signal peptide:**

amino acids 1-16

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

**FIGURE 136**

MASYLYGVLEFAVGLCAPIYCVSPANAPSAYPRPSSTKSTPASQVYSLNTDFAERLYRRLVLE  
TPSQNIFFSPVSVSTSLAMLSLGAHSVTKTQILQGLGFNLTHTPESAIHQGFQHLVHSLTVP  
SKDLTLKMGSALFVKKELQLQANFLGNVKRLYEAEVFSTDFSNPSIAQARINSHVKKKTQ GK  
VVDIIQGLDLLTAMVLVNHIFFKAKWEKPFHLEYTRKNFPFLVGEQVTVQVPMMHQKEQFAF  
GVDTELNCFVLQMDYKGDVAFFVLPSKGKMRQLEQALSARTLIKWSHSLQKRWIEVFIPRF  
SISASYNLETILPKMGIQNAFDKNADFSGIAKRDSLQVSKATHKAVLDVSEEGTEATAATTT  
KFIVRSKDGPSYFTVSFNRTFLMMITNKATDGILFLGKVENPTKS

**Signal peptide:**

amino acids 1-20

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250

[illegible]

**FIGURE 138**

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATI  
SGSSVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSE  
SSTPSSGASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSS  
TTSSGASTATNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTT  
SNGAGTATNSESSTTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSS  
GASTATNSDSSTTSSGAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGA  
NTATNSESSTVSSGASTATNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEAST  
ATNSESSTVSSGISTVTNSESSTTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAV  
SEAKPGGSLVPWEIFLITLVSVVAAVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGP  
GPGGNHGAPHRPRWSPNWFWRPVS SIAMEMSGRNSGP

**Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 510-532

**FIGURE 139**

GGGAGAGAGGATAAATAGCAGCGTGGCTTCCCTGGCTCCTCTCTGCATCCTTCCCGACCTTC  
CCAGCAAT**ATG**CATCTTGCACGTCTGGTCGGCTCCTGCTCCCTCCTTCTGCTACTGGGGGCC  
CTGTCTGGATGGGCGGCCAGCGATGACCCCATTGAGAAGGTCATTGAAGGGATCAACCGAGG  
GCTGAGCAATGCAGAGAGAGAGGTGGGCAAGGCCCTGGATGGCATCAACAGTGGAATCACGC  
ATGCCGGAAGGGAAGTGGAGAAGGTTTTCAACGGACTTAGCAACATGGGGAGCCACACCGGC  
AAGGAGTTGGACAAAGGCGTCCAGGGGCTCAACCACGGCATGGACAAGGTTGCCCATGAGAT  
CAACCATGGTATTGGACAAGCAGGAAAGGAAGCAGAGAAGCTTGGCCATGGGGTCAACAACG  
CTGCTGGACAGGCCGGAAGGAAGCAGACAAAGCGGTCCAAGGGTTCCACACTGGGGTCCAC  
CAGGCTGGGAAGGAAGCAGAGAACTTGGCCAAGGGGTCAACCATGCTGCTGACCAGGCTGG  
AAAGGAAGTGGAGAAGCTTGGCCAAGGTGCCCACCATGCTGCTGGCCAGGCCGGAAGGAGC  
TGCAGAATGCTCATAATGGGGTCAACCAAGCCAGCAAGGAGGCCAACCAGCTGCTGAATGGC  
AACCATCAAAGCGGATCTTCCAGCCATCAAGGAGGGGCCACAACCACGCCGTTAGCCTCTGG  
GGCCTCAGTCAACACGCCTTTCATCAACCTTCCCGCCCTGTGGAGGAGCGTCGCCAACATCA  
TGCCCT**TAA**ACTGGCATCCGGCCTTGCTGGGAGAATAATGTCGCCGTTGTCACATCAGCTGAC  
ATGACCTGGAGGGGTGGGGGTGGGGGACAGGTTTCTGAAATCCCTGAAGGGGGTGTACTG  
GGATTTGTGAATAAACTTGATACACCA



**FIGURE 140**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66675

><subunit 1 of 1, 247 aa, 1 stop

><MW: 25335, pI: 7.00, NX(S/T): 0

MHLARLVGSCSLLLLLLGALSGWAASDDPIEKVIEGINRGLSNAEREVGKALDGINSGITHAG  
REVEKVFNGLSNMGSHTGKELDKGVQGLNHGMDKVAHEINHGIGQAGKEAEKLGHGVNNAAG  
QAGKEADKAVQGFHTGVHQAGKEAEKLGQGVNHAADQAGKEVEKLGQGAHHAAGQAGKELQN  
AHNGVNQASKEANQLLNGNHQSGSSSHQGGATTTPLASGASVNTPFINLPALWRSVANIMP

**Important features of the protein:****Signal peptide:**

amino acids 1-25

**Homologous region to circumsporozoite (CS) repeats:**

amino acids 35-225

**FIGURE 141**

CTCCGGGTCCCCAGGGGCTGCGCCGGGCGGCCCTGGCAAGGGGGACGAGTCAGTGGACACTCCAGGAAGAGCGGC  
 CCCGCGGGGGCGATGACCGTGCGCTGACCGTGACTCACTCCAGGTCCGGAGGCGGGGGCCCCGGGGCGACTCG  
 GGGGCGGACCGGGGCGGAGCTGCCGCCCGTGAGTCCGGCCGAGCCACCTGAGCCCCGAGCCGCGGGACACCGTC  
 GCTCCTGCTCTCCGAATGCTGCGCACCGCGATGGGCCCTGAGGAGCTGGCTCGCCGCCCATGGGGCGCGTGCCG  
 CCTCGGCCACCGCTGCTGCTGCTCCTGCTGCTGCTGCTCCTGCTGCAGCCGCCGCTCCGACCTGGGCGCTCAGC  
 CCCCAGATCAGCCTGCCTCTGGGCTCTGAAGAGCGGCCATTCTCAGATTCTGAAGCTGAACACATCTCCAACACTAC  
 ACAGCCCTTCTGCTGAGCAGGGATGGCAGGACCCTGTACGTGGGTGCTCGAGAGGCCCTCTTTGCACTCAGTAGC  
 AACCTCAGCTTCTGCGCAGGCGGGGAGTACCAGGAGCTGCTTTGGGGTGACAGCGCAGAGAAGAAACAGCAGTGC  
 AGCTTCAAGGGCAAGGACCCACAGCGCGACTGTCAAACTACATCAAGATCCTCCTGCCGCTCAGCGGCAGTCAC  
 CTGTTCACTGTGGCACAGCAGCCTTCAAGCCCCATGTGTACCTACATCAACATGGAGAATTACCCCTGGCAAGG  
 GACGAGAAGGGGAATGTCTCCTGGAAGATGGCAAGGGCCGTTGTCCCTTCGACCCGAATTTCAAGTCCACTGCC  
 CTGGTGGTTGATGGCGAGCTCTACACTGGAACAGTCAGCAGCTTCCAAGGGAATGACCCGGCCATCTCGCGGAGC  
 CAAAGCCTTCGCCCCACCAAGACCGAGAGCTCCCTCAACTGGCTGCAAGACCCAGCTTTTGTGGCCTCAGCCTAC  
 ATTCTGAGAGCCTGGGCAGCTTGAAGGCGATGATGACAAGATCTACTTTTTCTCAGCGAGACTGGCCAGGAA  
 TTTGAGTTCTTTGAGAACACCATTGTGTCCCGCATTGCCCCATCTGCAAGGGCGATGAGGGTGGAGAGCGGGTG  
 CTACAGCAGCGCTGGACCTCCTTCTCAAGGCCAGCTGCTGTGCTCACGGCCCGACGATGGCTTCCCTTCAAC  
 GTGCTGCAGGATGTCTTACGCTGAGCCCCAGCCCCAGGACTGGCGTGACACCCCTTTTCTATGGGGTCTTCACT  
 TCCAGTGGCACAGGGGAATACAGAAGGCTCTGCCGTCTGTGCTTCAATGAAGGATGTGCAGAGAGTCTTC  
 AGCGGCTCTACAAGGAGGTGAACCGTGAGACACAGCAGTGGTACACCGTGACCCACCCGGTGCCACACCCCGG  
 CCTGGAGCGTGTCATCACCACAGTGCCCGGGAAGGAAGATCAACTCATCCTGCAGTCCCAGACCGCGTGCTG  
 AACTTCTCAAGGACCACTTCTGATGGACGGGCGAGTCCGAAGCCGATGCTGCTGCTGCAGCCCCAGGCTCGC  
 TACCAGCGCGTGGCTGTACACCGCGTCCCTGGCCTGCACCACACCTACGATGTCTCTTCTGGGCACTGGTGAC  
 GGCCGGCTCCACAAGGCAGTGAGCGTGGGCCCCCGGGTGACATCATTGAGGAGCTGCAGATCTTCTCATCGGGA  
 CAGCCCGTGAGAATCTGCTCCTGGACACCCACAGGGGGCTGCTGTATGCGGCCCTCACACTCGGGCGTAGTCCAG  
 GTGCCCATGGCCAACCTGCAGCCTGTACCGGAGCTGTGGGACTGCCTCCTCGCCCGGGACCCCTACTGTGCTTGG  
 AGCGGCTCCAGCTGCAAGCACGTGAGCCTCTACCAGCCTCAGCTGGCCACCAGGCGCTGGATCCAGGACATCGAG  
 GGAGCCAGCGCCAAGGACCTTTGAGCGCGCTTTCGGTTGTGTCCCGTCTTTTGTACCAACAGGGGAGAAGCCA  
 TGTGAGCAAGTCCAGTTCAGCCCAACACAGTGAACACTTTGGCCTGCCCGCTCCTCTCCAACCTGGCGACCCGA  
 CTCTGGCTACGCAACGGGGCCCCCGTCAATGCCTCGGCCCTCTGCCAGTGCTACCCACTGGGGACCTGCTGCTG  
 GTGGGCACCCAACAGCTGGGGGAGTTCCAGTGCTGGTCACTAGAGGAGGGCTTCCAGCAGCTGGTAGCCAGCTAC  
 TGCCAGAGGTGGTGGAGGACGGGGTGGCAGACCAACAGATGAGGGTGGCAGTGTACCCGTATTATCAGCACA  
 TCGCGTGTGAGTGACCAAGCAGCTGGTGGCAAGGCCAGCTGGGGTGACAGAGGTCTACTGGAAGGAGTTCCTGGTG  
 ATGTGCACGCTCTTTGTGCTGGCCGTGCTGCTCCAGTTTTATTCTTGTCTTACCGGCACCGGAACAGCATGAAA  
 GTCTTCTGAAGCAGGGGGAATGTGCCAGCGTGACCCCAAGACCTGCCCTGTGGTGCTGCCCCCTGAGACCCGC  
 CCACTCAACGGCCTAGGGCCCCCTAGCACCCCGCTCGATACCGAGGGTACCAGTCCCTGTGAGACAGCCCCCG  
 GGGGCCCCGAGTCTTCACTGAGTCAGAGAAGAGGCCACTCAGCATCCAAGACAGCTTCGTGGAGGTATCCCCAGTG  
 TGGCCCCGGCCCCGGGTCGCCCTTGGCTCGGAGATCCGTGACTCTGTGGTGTGAGAGCTGACTTCCAGAGGACGC  
 TGCCCTGGCTTCAGGGGCTGTGAATGCTCGGAGAGGGTCAACTGGACCTCCCTCCGCTCTGCTCTTCTGGAAC  
 ACGACCGTGGTGCCCGGCCCTTGGGAGCCTTGGAGCCAGCTGGCCTGCTGCTCTCCAGTCAAGTAGCGAAGCTCC  
 TACCACCCAGACACCCAACAGCCGTGGCCCCAGAGGTCTGGCCAAATATGGGGGCTGCCTAGGTTGGTGAA  
 CAGTGCTCCTTATGTAACTGAGCCCTTTGTTTTAAAAACAATTCCAAATGTGAACTAGAATGAGAGGGAAGAG  
 ATAGCATGGCATGCAGCACACAGGCTGCTCCAGTTTATGGCCTCCAGGGGTGCTGGGGATGCATCCAAAGTGG  
 TTGCTGAGACAGAGTTGGAACCCCTACCAACTGGCCTCTTCACTTCCACATTATCCCGCTGCCACCGGCTGC  
 CCTGTCTCACTGCAGATTCAAGGACAGCTTGGGCTGCGTGCGTTCTGCCTTGCCAGTCAGCCGAGGATGTAGTTG  
 TTGCTGCCGTGCTCCACACCTCAGGGACAGAGGGCTAGGTTGGCACTGCGGCCCTCACCAGGTCTGGGCTC  
 GGACCCAACCTCTGGACCTTTCCAGCCTGTATCAGGCTGTGGCCACACGAGAGGACAGCGCGAGCTCAGGAGAGA  
 TTTCTGTGACAATGTACGCCCTTCCCTCAGAATTCAGGGAAGAGACTGTGCGCTGCCTTCTCCGTTGTTGCGTGA  
 GAACCCGTGTGCCCTTCCCACCATATCCACCCTCGCTCCATCTTTGAACTCAAACACGAGGAACCTAAGTGCACC  
 CTGGTCTCTCCCGAGTCCCGAGTTCAACCTCCATCCCTCACCTTCTTCACTCTAAGGATATCAACACTGCCC  
 AGCACAGGGGCCCTGAATTTATGTGGTTTTTATACATTTTTTAATAAGATGCACTTTATGTATTTTTTAATAAA  
 GTCTGAAGAATTACTGTTTAAAAA

**FIGURE 142**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA67962

><subunit 1 of 1, 837 aa, 1 stop

><MW: 92750, pI: 7.04, NX(S/T): 6

MLRTAMGLRSWLAAPWGALPPRPPLLLLLLLLLLLLLLQPPPPTWALSPRISLPLGSEERPFLRF  
EAEHISNYTALLLSRDGRTLYVGAREALFALSSNLSFLPGGEYQELLWGADAEKKQQCSFKG  
KDPQRDCQNYIKILLPLSGSHLFTCGTAAAFSPMCTYINMENFTLARDEKGNVLLEDGKGRCP  
FDPNFKSTALVVDGELYTGTVSSFQGNDAISRSQSLRPTKTESSLNWLQDPAFVASAYIPE  
SLGSLQGDDDKIYFFFSETGQEFFEFFENTIVSRIARICKGDEGGERVLQQRWTSFLKAQLLC  
SRPDDGFPPFNVLQDVFTLSPSPQDWRDTLFYGVFTSQWHRGTTEGSAVCVFTMKDVQRVFSG  
LYKEVNRETQQWYTVTHPVPTPRPGACITNSARERKINSSQLPDRVLNFLKDHFLMDGQVR  
SRMLLLQPQARYQRVAVHRVPGLHHTYDVLFLGTGDGRLHKAVSVGPRVHIIIEELQIFSSGQ  
PVQNLLLDTHRGLLYAASHSGVVQVPMANCSLYRSCGDCLLARDPYCAWSGSSCKHVSlyQP  
QLATRPWIQDIEGASAKDLCSASSVVSPSFVPTGEKPCEQVQFQPNVTNLTACPLLSNLATR  
LWLRNGAPVNASASCHVLPTGDLLLVTQQLGEFQCWSLEEGFQQLVASYCPEVVEDGVADQ  
TDEGGSVPVIIISTSRVSAPAGGKASWGADRSYWKEFLVMCTLFVLAVLLPVLFLLYRHRNSM  
KVFLKQGECAVHPKTCPVVLPPETRPLNGLGPPSTPLDHRGYQSLSDSPPGARVFTSEKR  
PLSIQDSFVEVSPVCPRPRVRLGSEIRDSVV

**Transmembrane domains:**

amino acids 23-46 (type II), 718-738

**FIGURE 143A**

**FIGURE 143B**

TTAATTTTATATTCCTTACTGTTTTACTAAATATTAAGTGTTCTTTGACAATTTTGGTGCTCATGTGTTTTGGG  
GACAAAAGTGAAATGAATCTGTCAATTATACCAGAAAGTTAAATTCTCAGATCAAATGTGCCTTAATAAATTTGTT  
TTCATTTAGATTTCAAACAGTGATAGACTTGCCATTTTAATACACGTCATTGGAGGGCTGCGTATTTGTAAATAG  
CCTGATGCTCATTGGGAAAAATAAACAGTGAACAATATTTTTCTATTGTACTTTTCGAACCATTTTGTCTCATT  
ATTCCTGTTTTAGCTGAAGAATTGTATTACATTTGGAGAGTAAAAACTTAAACACGAAAAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 144**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68836

><subunit 1 of 1, 802 aa, 1 stop

><MW: 91812, pI: 9.52, NX(S/T): 3

MAARGRRRAWLSVLLGLVLGFLASRLVLPRASELKRAGPRRRASPEGCRSGQAAASQAGGAR  
GDARGAQLWPPGSDPDGGPRDRNFLVGVMTAQKYLQTRAVAAYRTWSKTIPGKVQFFSSEG  
SDTSVPIPVVPLRGVDDSYPPQKKSFMMLKYMHDHYLDKYEFMRADDDVYIKGDRLENFLR  
SLNSSEPLFLGQTGLGTTEEMGKLALEPGENFCMGGPGVIMSREVLRRMVPHIGHKCLREMYT  
THEDVEVGRCVRRFAGVQCWWSYEMRQLFYENYEQNKKGYIRDLHNSKIHQAITLHPNKNPP  
YQYRLHSYMLSRKISELRHRTIQLHREIVLMSKYSNTEIHKEDLQLGIPPSFMRFPQPRQREE  
ILEWEFLTGKYLYSAVDGQPPRRGMDSAQREALDDIVMQVMEMINANAKTRGRIIDFKEIQY  
GYRRVNPYGAEYILDLLLLYKKHKGKKMTVPVRRHAYLQQTFSKIQFVEHEELDAQELAKR  
INQESGSLSFLSNSLKKLVFPQLPGSKSEHKEPKDKKINILIPLSGRFDMFVRFMGNFEKTC  
LIPNQNVKLVVLLFNSDSNPDKAKQVELMRDYRIKYPKADMQILPVSGEFSRALALEVGSSQ  
FNNESLLFFCDVDLVFETTEFLQRCRANTVLGQQIYFPIIFSQYDPKIVYSGKVPSDNHFAFT  
QKTGFWRNYGFGITCIYKGD LVRVGGFDVSIQGWGLEDDVLFNKVVQAGLKTFRSQEVGVVH  
VHHPVFCDPNLDPKQYKMCLGSKASTYGSTQQLAEMWLEKNDPSYSKSSNNNGSVRTA

**Signal peptide:**

amino acids 1-23

**FIGURE 145**

GGACAACCGTTGCTGGGTGTCCCAGGGCCTGAGGCAGGACGGTACTCCGCTGACACCTTCCC  
 TTTCGGCCTTGAGGTTCCCAGCCTGGTGGCCCCAGGACGTTCCGGTCGCATGGCAGAGTGCT  
 ACGGACGACGCCT**ATG**AAGCCCTTAGTCCTTCTAGTTGCGCTTTTGCTATGGCCTTCGTCTG  
 TGCCGGCTTATCCGAGCATAACTGTGACACCTGATGAAGAGCAAACTTGAATCATTATATA  
 CAAGTTTTAGAGAACCTAGTACGAAGTGTTCCCTCTGGGGAGCCAGGTCGTGAGAAAAAATC  
 TAACTCTCCAAACATGTTTATTCTATAGCATCAAAGGGATCAAATTTAAGGAGCTAGTTA  
 CACATGGAGACGCTTCAACTGAGAATGATGTTTTAACCAATCCTATCAGTGAAGAACTACA  
 ACTTTCCTTACAGGAGGCTTCACACCGGAAATAGGAAAGAAAAACACACGGAAAGTACCCC  
 ATTCTGGTCGATCAAACCAAACAATGTTTCCATTGTTTTGCATGCAGAGGAACCTTATATTG  
 AAAATGAAGAGCCAGAGCCAGAGCCGGAGCCAGCTGCAAAACAACTGAGGCACCAAGAATG  
 TTGCCAGTTGTTACTGAATCATCTACAAGTCCATATGTTACCTCATACAAGTCACCTGTCAC  
 CACTTTAGATAAGAGCACTGGCATTGAGATCTCTACAGAATCAGAAGATGTTCCCTCAGCTCT  
 CAGGTGAACTGCGATAGAAAAACCCGAAGAGTTTGGAAGCAGGAGAGTTGGAATAAT  
 GATGACATTTTGAAAAAATTTTAGATATTAATTCACAAGTGCAACAGGCACTTCTTAGTGA  
 CACCAGCAACCCAGCATATAGAGAAGATATTGAAGCCTCTAAAGATCACCTAAAACGAAGCC  
 TTGCTCTAGCAGCAGCAGCAGAACATAAATTAATAACAATGTATAAGTCCCAGTTATTGCCA  
 GTAGGACGAACAAGTAATAAAATTGATGACATCGAACTGTTATTAAACATGCTGTGTAATTC  
 TAGATCTAACTCTATGAATATTTAGATATTAATGTGTTCCACCAGAGATGAGAGAAAAAG  
 CTGCTACAGTATTCAATACATTAATAAATATGTGTAGATCAAGGAGAGTCACAGCCTTATTA  
 AAAGTTTAT**TAA**ACAATAATATAAAAAATTTTAAACCTACTTGATATTCCATAACAAAGCTGA  
 TTTAAGCAAACCTGCATTTTTTTCACAGGAGAAATAATCATATTTCGTAATTTCAAAGTTGTAT  
 .AAAAATATTTTCTATTGTAGTTCAAATGTGCCAACATCTTTATGTGTCATGTGTTATGAACA  
 ATTTTCATATGCACTAAAAACCTAATTTAAATAAAAAATTTTGGTTCAGGAAAAA

### FIGURE 146

MKPLVLLVALLLWPSSVPAYPSITVTPDEEQNLNHYIQVLENLVRVSPSGEPGREKKSNSPK  
 HVYSIASKGSKFKELVTHGDASTENDVLTNPISEETTTFTPTGGFTPEIGKKKHTSTPFWSI  
 KPNNVSIVLHAEOPYIENEEPEPEPEPAAKQTEAPRMLPVVTESSTSPYVTSYKSPVTTLDK  
 STGIEISTESEDVPQLSGETAIEKPEEFGKHPESWNNDDILKKILDINSQVQQALLSDTSNP  
 AYREDIEASKDHLKRSLALAAAAEHKLKTMYSQLLPVGRTSNKIDDIETVINMLCNSRSKL  
 YEYLDIKCVPPEMREKAATVFNTLKNMCRSRRVTALLKVY

amino acids 1-19



**FIGURE 147**

401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTG  
 ACCTCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTCATAGTGTGAGATCAACCCACA  
 GGAATATCC**ATG**GGCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACA  
 GTGGCAAGTCACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCT  
 CCTGCTCCCTCTTTCTGAGACCAGTGCAGAGGCTATGGAAGTGC GGTTCTTCAGGAATCAG  
 TTCCATGCTGTGGTCCACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACA  
 GTATCGAGGGAGAACTGAGTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTCTAAGGC  
 TAAAAACATCACTCCCTCGGACATCGGCCTGTATGGGTGCTGGTTCAGTTCCCAGATTTAC  
 GATGAGGAGGCCACCTGGGAGCTGCGGGTGGCAGCACTGGGCTCACTTCCTCTCATTTCCAT  
 CGTGGGATATGTTGACGGAGGTATCCAGTTACTCTGCCTGTCCTCAGGCTGGTTCCTCCAGC  
 CCACAGCCAAGTGGAAAGGTCCACAAGGACAGGATTTGTCTTCAGACTCCAGAGCAAATGCA  
 GATGGGTACAGCCTGTATGATGTGGAGATCTCCATTATAGTCCAGGAAAATGCTGGGAGCAT  
 ATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGAGGTGGAATCCAAGGTATTGATAGGAG  
 AGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTTCTATTTTACTCGGGTTACTCTGTGGT  
 GCCCTGTGTGGTGTGTGTCATGGGGATGATAATTGTTTTCTTCAAATCCAAAGGGAAAAATCCA  
 GGCGGAACTGGACTGGAGAAGAAAGCACGGACAGGCAGAAATTGAGAGACGCCCGGAAACACG  
 CAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCGTTTTCTGATCTGAAA  
 ACTGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAAGAGATTTACAAGGAA  
 GAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGACGTGGGAC  
 AAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGGAAGAACAATGTG  
 ACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTTAC  
 ATTCAATCCCCATTTTATCAGCCTCCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCC  
 TGGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACC  
 CTGCTGACATGTCAGTTTGAAGGCTTGTTGAGACCCTATATCCAGCATGCGATGTATGACGA  
 GGAAAAGGGGACTCCCATATTATATGTCCAGTGTCTGGGGAT**TGAG**ACAGAGAAGACCCTG  
 CTTAAAGGGCCCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGCTCCCGACAGGTGGC  
 CCCAGCTTCCTCTCCGGAGCCTGCGCACAGAGAGTCACGCCCCCACTCTCCTTTAGGGAGC  
 TGAGGTTCTTCTGCCCTGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGAT  
 TGGCCTGACCCTGTGGGAGTCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCA  
 CATTAGGTTTtagTTTGTGAAAACCTCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCC  
 CAGGCTCCTCATTTGCTAGTCACGGACAGTGATTCTTGCCTCACAGGTGAAGATTAAAGAGA  
 CAACGAATGTGAATCATGCTTGCAGGTTTGAGGGCACAGTGTGTTGCTAATGATGTGTTTTTA  
 TATTATACATTTTCCCACCATAAACTCTGTTTGCTTATTTCCACATTAATTTACTTTTCTCTA  
 TACCAAATCACCCATGGAATAGTTATTGAACACCTGCTTTGTGAGGCTCAAAGAATAAAGAG  
 GAGGTAGGATTTTCACTGATTCTATAAGCCCAGCATTACCTGATACCAAACCAGGCAAAG  
 AAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCCCTCATTAACACAGACACAAAAA  
 TTCTAAATAAAATTTTAAACAATTAACCTAAACAATATTTTAAAGATGATATATAACTACT  
 CAGTGTGGTTTGTCCCACAAATGCAGAGTTGGTTTAAATATTTAAATATCAACCAGTGTAATT  
 CAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAAAAAAAAAAA

**FIGURE 148**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68866

><subunit 1 of 1, 466 aa, 1 stop

><MW: 52279, pI: 6.16, NX(S/T): 2

MAFVLILVLSFYELVSGWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHA  
VVHLYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEE  
ATWELRVAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGY  
SLYDVEISIIIVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALC  
GVVMGMIIVFFKSKGKIQAELDWRRKHGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVT  
HRKAPQEVPHSEKRFTRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLS  
PNNGYWVLRLTTEHLYFTFNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLT  
CQFEGLLRPYIQHAMYDEEKGTPIFICPVSWG

**Signal peptide:**

amino acids 1-17

**Transmembrane domains:**

amino acids 131-150, 235-259

**FIGURE 149**

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTA  
 GGAAAAGAGTTTGTGGGAACCCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTC  
 CTGGCAGTGTGCATTGGACTCACTGTTTATTATGTGAGATATAATCAAAAGAAGACCTACAA  
 TTACTATAGCACATTGTCAATTTACAACCTGACAACTATATGCTGAGTTTGGCAGAGAGGCTT  
 CTAACAATTTTACAGAAATGAGCCAGAGACTTGAATCAATGGTGAAAAATGCATTTTATAAA  
 TCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTCAGTCAACAGAAGCATGG  
 AGTGTTGGCTCATATGCTGTTGATTTGTAGATTTCACTCTACTGAGGATCCTGAAACTGTAG  
 ATAAATTTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTA  
 GATCCTCACTCAGTTAAAATTAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCA  
 TTGCTGCGGAACACGAAGAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAG  
 AAGTAGAAGAGGGTGAATGGCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGT  
 GGAGCAACCTTAATTAATGCCACATGGCTTGTGAGTGCTGCTCACTGTTTTACAACATATAA  
 GAACCTGCCAGATGGACTGCTTCCTTTGGAGTAACAATAAACCTTCGAAAATGAAACGGG  
 GTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACATGACTATGATATTTCT  
 CTTGCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCTCCCTGA  
 TGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAA  
 ATGATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAAC  
 TGCAATGAACCTCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTT  
 AGAAGGAAAAACAGATGCATGCCAGGGTGACTCTGGAGGACCACTGGTTAGTTCAGATGCTA  
 GAGATATCTGGTACCTTGCTGGAATAGTGAGCTGGGGAGATGAATGTGCGAAACCCAACAAG  
 CCTGGTGTTTATACTAGAGTTACGGCCTTGCGGGACTGGATTACTTCAAAAACCTGGTATCTA  
 AGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTTGTTTTTTGGGTGTGGAGGCCATT  
 TTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCAATAAAC  
 TGTTTGCTTGATGCATGTATTTTCTTCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTG  
 CCAGATCAACTCTGTCATCTGTGAGCAATAGTTGAACTTTATGTACATAGAGAAATAGATA  
 ATACAATATTACATTACAGCCTGTATTCAATTTGTTCTCTAGAAGTTTTGTGAGAATTTTGAC  
 TTGTTGACATAAATTTGTAATGCATATATAACAATTTGAAGCACTCCTTTTCTTCAGTTCCTC  
 AGCTCCTCTCATTTTCAGCAAATATCCATTTTCAAGGTGCAGAACAAGGAGTGAAAGAAAATA  
 TAAGAAGAAAAAATCCCCTACATTTTATTGGCACAGAAAAGTATTAGGTGTTTTTCTTAGT  
 GGAATATTAGAAATGATCATATTCATTATGAAAGGTCAAGCAAAGACAGCAGAATACCAATC  
 ACTTCATCATTTAGGAAGTATGGGAACCTAAGTTAAGGAAGTCCAGAAAGAAGCCAAGATATA  
 TCCTTATTTTTCATTTCCAAACAACCTACTATGATAAATGTGAAGAAGATTCTGTTTTTTTGTG  
 ACCTATAATAATTATACAACTTCATGCAATGTACTTGTCTAAGCAAATTAAGCAAATAT  
 TTATTTAACATTGTTACTGAGGATGTCAACATATAACAATAAAATATAAATCACCCA

**FIGURE 150**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68871

><subunit 1 of 1, 423 aa, 1 stop

><MW: 47696, pI: 8.96, NX(S/T): 3

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTD  
KLYAEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICR  
FHSTEDPETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTL  
GQSLRIVGGTEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFG  
VTIKPSKMKRGLRRIIVHEKYKHPSHDYDISLAELSSVPYTNVHRVCLPDASYEFQPGDV  
MFVTGFGALKNDGYSONHLRQAQVTLLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGD  
SGGPLVSSDARDIWIYLAGIVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

**Transmembrane domain:**

amino acids 21-40 (type II)

**FIGURE 151**T  
o  
t  
a  
l  
c  
o  
u  
n  
t  
=  
6  
9  
3  
4  
4

GTCGAAGGTTATAAAAAGCTTCCAGCCAAACGGCATTGAAGTTGAAGATACAACCTGACAGCA  
CAGCCTGAGATCTTGGGGATCCCTCAGCCTAACACCCACAGACGTCAGCTGGTGGATTCCCC  
CTGCATCAAGGCCTACCCACTGTCTCCATGCTGGGCTCTCCCTGCCTTCTGTGGCTCCTGGC  
CGTGACCTTCTTGGTTCCCAGAGCTCAGCCCTTGGCCCCCTCAAGACTTTGAAGAAGAGGAGG  
CAGATGAGACTGAGACGGCGTGGCCGCCTTTGCCGGCTGTCCCCTGCGACTACGACCACTGC  
CGACACCTGCAGGTGCCCTGCAAGGAGCTACAGAGGGTCGGGCCGGCGGCCTGCCTGTGCCC  
AGGACTCTCCAGCCCCGCCAGCCGCCGACCCGCCGCGCATGGGAGAAGTGCGCATTGCGG  
CCGAAGAGGGCCGCGCAGTGGTCCACTGGTGTGCCCCCTTCTCCCCGGTCCCTCCACTACTGG  
CTGCTGCTTTGGGACGGCAGCGAGGCTGCGCAGAAGGGGGCCCCCGCTGAACGCTACGGTCCG  
CAGAGCCGAACGAAGGGGGCTGAAGCCAGGGGGCATTTATGTCGTTTGCGTAGTGGCCGCTA  
ACGAGGCCGGGGCAAGCCGCGTGCCCCAGGCTGGAGGAGAGGGCCTCGAGGGGGCCGACATC  
CCTGCCTTCGGGCCTTGACAGCCGCCTTGCGGTGCCGCCCAACCCCCGCACTCTGGTCCACGC  
GGCCGTCGGGGTGGGCACGGCCCTGGCCCTGCTAAGCTGTGCCGCCCTGGTGTGGCACTTCT  
GCCTGCGCGATCGCTGGGGCTGCCCCGCGCCGAGCCGCCGCCGAGCCGCAGGGGCGCTCTGA  
AAGGGGCCTGGGGGCATCTCGGGCACAGACAGCCCCACCTGGGGCGCTCAGCCTGGCCCCCG  
GGAAAGAGGAAAACCCGCTGCCTCCAGGGAGGGCTGGACGGCGAGCTGGGAGCCAGCCCCAG  
GCTCCAGGGCCACGGCGGAGTCATGGTTCTCAGGACTGAGCGCTTGTTTAGGTCCGGTACTT  
GGCGCTTTGTTTCCTGGCTGAGGTCTGGGAAGGAATAGAAAGGGGCCCCCAATTTTTTTTAA  
AGCGGCCAGATAATAAATAATGTAACCTTTGCGGTTAAAAAAAAAAAAAAAAAAAAA

**FIGURE 152**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68874

><subunit 1 of 1, 238 aa, 1 stop

><MW: 25262, pI: 6.44, NX(S/T): 1

MLGSPCLLWLLAVTFLVPRAQPLAPQDFEEEEADETETAWPPLPAVPCDYDHCRLQVPCKE  
LQRVGPAACLCPGLSSPAQPPDPFRMGVRIAAEEGRAVVHWCAPFSPVLHYWLLLWDGSEA  
AQKGPPPLNATVRRRAELKGLKPGGIYVVCVVAANEAGASRVPAQAGGEGLEGADIPAFGPCSRL  
AVPPNPRTLVAHAVGVGTALALLSCAALVWHFCLRDRWGCPRAAARAAGAL

**Important features of the protein:****Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 194-220

**N-glycosylation site.**

amino acids 132-135



**FIGURE 154**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68880

><subunit 1 of 1, 305 aa, 1 stop

><MW: 35383, pI: 5.99, NX(S/T): 0

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPV  
VHPVMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQ  
WSDMVTLKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMDWPPDSCCVREF  
PGCSKQAHQEDLSLDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWAL  
YYDRREPGTDQMMSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

**Signal peptide:**

amino acids 1-33

**Transmembrane domains:**

amino acids 12-35, 57-86, 94-114, 226-248



**FIGURE 155**

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCC  
 TGCCTCGGGCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGA  
 CCTGTGTGGGGAGGCCCTCCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAG  
 ACCGGGAGGATCACAGAGCCAGC**ATG**TTACAGGATCCTGACAGTGATCAACCTCTGAACAGC  
 CTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCCATGGAGACCTTCAGAAAGGTGGGGAT  
 CCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGTGGTTGTCCTCATCAAGG  
 TGATTCTGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGAGGAAGCAG  
 CTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTT  
 CCCCGAAGGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGG  
 ACTCGGCCACAGGGAAGTGGTTCTCTGCCTGTTTTGACAACCTTCACAGAAGCTCTCGCTGAG  
 ACAGCCTGTAGGCAGATGGGCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGA  
 TGTTGTTGAAATCACAGAAAACAGCCAGGAGCTTCGCATGCGGAACTCAAGTGGGCCCTGTC  
 TCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCTGTGGGAAGAGCCTGAAGACCCCCCGT  
 GTGGTGGGTGGGGAGGAGGCCTCTGTGGATTCTTGCCCTTGGCAGGTCAGCATCCAGTACGA  
 CAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCAGCCCACT  
 GCTTCAGGAAACATAACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGGC  
 AGCTTCCCATCCCTGGCTGTGGCCAAGATCATCATCATTGAATTCAACCCCATGTACCCCAA  
 AGACAATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGC  
 CCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGA  
 TGGGGCTTTACGAAGCAGAATGGAGGGGAAGATGTCTGACATACTGCTGCAGGCGTCAGTCCA  
 GGTCATTGACAGCACACGGTGCAATGCAGACGATGCGTACCAGGGGGGAAGTCACCGAGAAGA  
 TGATGTGTGCAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGTGACAGTGGTGGGCCC  
 CTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGGCTGCGG  
 GGGCCCCGAGCACCCCAAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATG  
 TCTGGAAGGCTGAGCTG**TAA**TGCTGCTGCCCCCTTTGCAGTGCTGGGAGCCGCTTCCTTCCTG  
 CCCTGCCCACCTGGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCCTGGGTACACCC  
 CTCTGCCCACAGCCTCAGCATTCTTGGAGCAGCAAAGGGCCTCAATTCCTGTAAGAGACCC  
 TCGCAGCCCAGAGGCGCCAGAGGAAGTCAGCAGCCCTAGCTCGGCCACACTTGGTGCTCCC  
 AGCATCCCAGGGAGAGACACAGCCCACTGAACAAGGTCTCAGGGGTATTGCTAAGCCAAGAA  
 GGAACCTTCCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAGGCCAGATCACTGTGGG  
 CTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAAGCCTA  
 CTAGAGCAAGAAACCAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTT  
 ACCTACTGTTGTCATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCT  
 CTGGCAAAAAAAAAAAAA

**FIGURE 156**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68885

><subunit 1 of 1, 432 aa, 1 stop

><MW: 47644, pI: 5.18, NX(S/T): 2

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVVVLIVILDKYYF  
LCGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWF  
SACFDNFTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSL  
HCLACGKSLKTPRVVGGEASVDSWPWQVSIQYDKQHVCSSILDPHWVLTAHCERKHTDV  
FNWKVRAGSDKLGSPSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFD  
EELTPATPLWIIIGWGFQNGGKMSDILLQASVQVIDSTRCNADDAAYQGEVTEKMMCAGIPE  
GGVDTTCQGDSSGGLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

**Transmembrane domain:**

amino acids 32-53 (typeII)

**FIGURE 157**

GGGCTGAGGCACTGAGAGACCGGAAAGCCTGGCATTCCAGAGGGAGGGAAACGCAGCGGCATCCCCAGGCTCCAG  
 AGCTCCCTGGTGACAGTCTGTGGCTGAGCATGGCCCTCCCAGCCCTGGGCCTGGACCCCTGGAGCCTCCTGGGCC  
 TTTTCCTCTTCCAACCTGCTTCAGCTGCTGCTGCCGACGACGACCGCGGGGGAGGCGGGCAGGGGGCCATGCCCCA  
 GGGTCAGATACTATGCAGGGGATGAACGTAGGGCACTTAGCTTCTTCCACCAGAAGGGCCTCCAGGATTTTGACA  
 CTCTGCTCCTGAGTGGTGATGGAAATACTCTCTACGTGGGGGCTCGAGAAGCCATTCTGGCCTTGGATATCCAGG  
 ATCCAGGGGTCCCCAGGCTAAAGAACATGATACCGTGGCCAGCCAGTGACAGAAAAAGAGTGAATGTGCCTTTA  
 AGAAGAAGAGCAATGAGACACAGTGTTCACCTTCATCCGTGTCTGGTTTCTTACAATGTCACCCATCTCTACA  
 CCTGCGGCACCTTCGCCTTCAGCCCTGCTTGTACCTTCATTGAACCTCAAGATTCCCTACCTGTTGCCATCTCGG  
 AGGACAAGGTCATGGAGGGAAAAGGCCAAAGCCCCTTTGACCCCGCTCACAAGCATACGGCTGTCTTGGTGGATG  
 GGATGCTCTATTCTGGTACTATGAACAACCTCCTGGGCAGTGAGCCCATCCTGATGCGCACACTGGGATCCCAGC  
 CTGTCTCAAGACCGACAACCTCCTCCGCTGGCTGCATCATGACGCCCTCCTTTGTGGCAGCCATCCCTTCGACCC  
 AGGTGCTCTACTTCTTCTTCGAGGAGACAGCCAGCGAGTTTGACTTCTTTGAGAGGCTCCACACATCGCGGGTGG  
 CTAGAGTCTGCAAGAATGACGTGGGCGGCGAAAAGCTGCTGCAGAAGAAGTGGACCACCTTCCTGAAGGGCCAGC  
 TGCTCTGCACCCAGCCGGGGCAGCTGCCCTTCAACGTCACTCCGCCACGCGGTCTGCTCCCCGCCGATTCTCCCCA  
 CAGCTCCCCACATCTACGCAGTCTTCACCTCCCAGTGGCAGGTTGGCGGGACCAGGAGCTCTGCGGTTTGTGCCT  
 TCTCTCTCTTGGACATTGAACGTGTCTTTAAGGGGAAATACAAAGAGTTGAACAAAGAACTTCACGCTGGACTA  
 CTTATAGGGGCCCTGAGACCAACCCCGGCCAGGCAGTTGCTCAGTGGGCCCTCCTCTGATAAGGCCCTGACCT  
 TCATGAAGGACCATTTCTGATGGATGAGCAAGTGGTGGGGACGCCCCCTGCTGGTGAAATCTGGCGTGGAGTATA  
 CACGGCTTGAGTGGAGACAGCCAGGGCCTTGATGGGCACAGCCATCTTGTGATGACCTGGGAACCAACACAG  
 GGTGCTCCACAAGGCTGTGGTAAGTGGGGACAGCAGTGCTCATCTGGTGGAAGAGATTGAGCTGTTCCCTGACC  
 CTGAACCTGTTGCAACCTGCAGCTGGCCCCACCCAGGGTGCAGTGTGTGAGGCTTCTCAGGAGGTGTCTGGA  
 GGGTGCCCCGAGCCAACCTGTAGTGTCTATGAGAGCTGTGTGGACTGTGTCTTGGCCGGGACCCCCACTGTGCCT  
 GGGACCCTGAGTCCCGAACCTGTTGCCTCCTGTCTGCCCCAACCTGAACTCCTGGAAGCAGGACATGGAGCGGG  
 GGAACCCAGAGTGGGCATGTGCCAGTGGCCCCATGAGCAGGAGCCTTCGGCCTCAGAGCCGCCCCGAAATCATT  
 AAGAAGTCTGGCTGTCCCCAACTCCATCCTGGAGCTCCCCTGCCCCACCTGTCAGCCTTGGCCTCTTATTATT  
 GGAGTCATGGCCCAGCAGCAGTCCCAGAAGCCTCTTCCACTGTCTACAATGGCTCCCTCTTGCTGATAGTGCAGG  
 ATGGAGTTGGGGTCTCTACAGTGTGGGCAACTGAGAATGGCTTTTCATACCCTGTGATCTCCTACTGGGTGG  
 ACAGCCAGGACCAGACCCTGGCCCTGGATCCTGAACTGGCAGGCATCCCCGGGAGCATGTGAAGGTCCCGTTGA  
 CCAGGGTCACTGGTGGGGCCGCCCTGGCTGCCAGCAGTCTACTGGCCCCACTTTGTCACTGTCACTGTCTCTCT  
 TTGCCTTAGTGTCTTTCAGGAGCCCTCATCATCCTCGTGGCCTCCCCATTGAGAGCACTCCGGGCTCGGGGCAAGG  
 TTCAGGGCTGTGAGACCCTGCGCCCTGGGGAGAAGGCCCGTTAAGCAGAGAGCAACACCTCCAGTCTCCCAAGG  
 AATGCAGGACCTCTGCCAGTGATGTGGACGCTGACAACAACCTGCCTAGGCACTGAGGTAGCTTAAACTCTAGGCA  
 CAGGCCGGGGCTGCGGTGCAGGCACCTGGCCATGCTGGCTGGGCGGCCCAAGCACAGCCCTGACTAGGATGACAG  
 CAGCACAAAAGACCACCTTTCTCCCCTGAGAGGAGCTTCTGCTACTCTGCATCACTGATGACACTCAGCAGGGTG  
 ATGCACAGCAGTCTGCCTCCCCTATGGGACTCCCTTCTACCAAGCACATGAGCTCTCTAACAGGGTGGGGGCTAC  
 CCCCAGACCTGCTCCTACACTGATATTGAAGAACCTGGAGAGGATCCTTCAGTTCTGGCCATTCCAGGGACCCTC  
 CAGAAACACAGTGTTCAGAGACCCCTAAAAAACCTGCCTGTCCAGGACCCTATGGTAATGAACACCAACATC  
 TAAACAATCATATGCTAACATGCCACTCCTGGAACTCCACTCTGAAGCTGCCGCTTTGGACACCAACACTCCCT  
 TCTCCCAGGGTCATGCAGGGATCTGCTCCCTCCTGCTTCCCTTACCAGTCGTGCACCGCTGACTCCCAGGAAGTC  
 TTTCTGAAGTCTGACCACCTTTCTTCTTGCTTCAGTTGGGGCAGACTCTGATCCCTTCTGCCCTGGCAGAATGG  
 CAGGGGTAATCTGAGCCTTCTTCACTCCTTTACCCTAGCTGACCCCTTCACCTCTCCCCCTCCCTTTTCTTTGT  
 TTTGGGATTGAGAAAACCTGCTTGTGAGAGACTGTTTATTTTTTATTAAAAATATAAGGCTTAAAAAAA

**FIGURE 158**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71166

><subunit 1 of 1, 761 aa, 1 stop

><MW: 83574, pI: 6.78, NX(S/T): 4

MALPALGLDPWSLLGLFLFQLLQLLLPTTTAGGGGQGMPRVRYAGDERRALSFFHQKGLQ  
DFDTLLLSGDGNTLYVGAREAILALDIQDPGVPRLNMIWPASDRKKSECAFKKKSNETQC  
FNFIRVLVSYNVTHLYTCGTFAFSPACTFIELQDSYLLPISEDKVMEGKGQSPFDPAHKHTA  
VLVDGMLYSGTMNNFLGSEPILMRTLGSQPVLKTDNFLRWLHHDASFVAAIPSTQVVYFFFE  
ETASEFDFFERLHTSRVARVCKNDVGGEKLLQKKWTTFLKAQLLCTQPGQLPFNVIRHAVLL  
PADSPTAPHIYAVFTSQWQVGGTRSSAVCAFSLLDIERVFKGKYKELNKETSRWTTYRGPET  
NPRPGSCSVGPSSDKALTFMKDHFLMDEQVVGTPLLVKSGVEYTRLAVETAQGLDGHSHLVM  
YLGTTTGSLSHKAVVSGDSSAHLVEEIQLFDPPEPVRNLQLAPTQGAVFVGFSGGVWRVPRAN  
CSVYESCVCVLDARPHCAWDPESTRCCLLSAPNLSWKQDMERGNPEWACASGPMSRSLRP  
QSRPQIIKEVLAVPNSILELPCPHLSALASYWSHGPAAVPEASSTVYNGSLLLIVQDGVGG  
LYQCWATENGFSYPVISYWVDSQDQTLALDPELAGIPREHVKVPLTRVSGGAALAAQQSYWP  
HFVTVTVLFLVLGALIILVASPLRALRARGKVQGCETLRPGEKAPLSREQHLQSPKECRT  
SASDVDADNNCLGTEVA

**Signal peptide:**

amino acids 1-30

**Transmembrane domains:**

amino acids 136-156, 222-247, 474-490, 685-704

**FIGURE 159**

AGGGTCCCTTAGCCGGGCGCAGGGCGCGCAGCCCAGGCTGAGATCCGCGGCTTCCGTAGAAG  
 TGAGC**ATG**GCTGGGCAGCGAGTGCTTCTTCTAGTGGGCTTCCTTCTCCCTGGGGTCCTGCTC  
 TCAGAGGCTGCCAAAATCCTGACAATATCTACAGTAGGTGGAAGCCATTATCTACTGATGGA  
 CCGGGTTTCTCAGATTCTTCAAGATCACGGTCATAATGTCACCATGCTTAACCACAAAAGAG  
 GTCCTTTTATGCCAGATTTTAAAAAGGAAGAAAAATCATATCAAGTTATCAGTTGGCTTGCA  
 CCTGAAGATCATCAAAGAGAATTTAAAAAGAGTTTTGATTTCTTTCTGGAAGAACTTTAGG  
 TGGCAGAGGAAAATTTGAAAATTTAATGTTCTAGAATACTTGGCGTTGCAGTGCAGTC  
 ATTTTTTAAATAGAAAGGATATCATGGATTCTTAAAGAATGAGAACTTCGACATGGTGATA  
 GTTGAACTTTTGACTACTGTCCTTTCTGATTGCTGAGAAGCTTGGGAAGCCATTTGTGGC  
 CATTCTTTCCACTTCATTGCGCTCTTTGGAATTTGGGCTACCAATCCCCTTGTCTTATGTTT  
 CAGTATTCGCTTCCTTGCTGACTGATCACATGGACTTCTGGGGCCGAGTGAAGAATTTCTG  
 ATGTTCTTTAGTTTCTGCAGGAGGCAACAGCACATGCAGTCTACATTTGACAACACCATCAA  
 GGAACATTTACAGAAGGCTCTAGGCCAGTTTTGTCTCATCTTCTACTGAAAGCAGAGTTGT  
 GGTTCATTAACCTGACTTTGCCTTTGATTTTGCTCGACCTCTGCTTCCCAACACTGTTTAT  
 GTTGGAGGCTTGATGGAAAAACCTATTAAACCAGTACCACAAGACTTGGAGAACTTCATTGC  
 CAAGTTTGGGGACTCTGGTTTTGTCCTTGTGACCTTGGGCTCCATGGTGAACACCTGTCAGA  
 ATCCGGAAATCTTCAAGGAGATGAACAATGCCTTTGCTCACCTACCCCAAGGGGTGATATGG  
 AAGTGTCAGTGTTCTCATTGGCCCAAAGATGTCCACCTGGCTGCAAATGTGAAAATTTGTGGA  
 CTGGCTTCCTCAGAGTGACCTCCTGGCTCACCCAAGCATCCGTCTGTTTGTACCCACGGCG  
 GGCAGAAATAGCATAATGGAGGCCATCCAGCATGGTGTGCCCATGGTGGGGATCCCTCTCTTT  
 GGAGACCAGCCTGAAAACATGGTCCGAGTAGAAGCCAAAAAGTTTGGTGTCTTATTAGTT  
 AAAGAAGCTCAAGGCAGAGACATTGGCTCTTAAGATGAAACAAATCATGGAAGACAAGAGAT  
 ACAAGTCCGCGGCAGTGGCTGCCAGTGTATCCTGCGCTCCCACCCGCTCAGCCCCACACAG  
 CGGCTGGTGGGCTGGATTGACCACGTCTCCAGACAGGGGGCGCGACGCACCTCAAGCCCTA  
 TGTCTTTCAGCAGCCCTGGCATGAGCAGTACCTGTTGACGTTTTTGTGTTTCTGCTGGGGC  
 TCACTCTGGGGACTCTATGGCTTTGTGGGAAGCTGCTGGGCATGGCTGTCTGGTGGCTGCGT  
 GGGGCCAGAAAGGTGAAGGAGACAT**TAA**GGCCAGGTGCAGCCTTGGCGGGGTCTGTTTGGTGG  
 GCGATGTCACCATTTCTAGGGAGCTTCCCACTAGTTCTGGCAGCCCCATTCTCTAGTCCTTC  
 TAGTTATCTCCTGTTTTCTTGAAGAACAGGAAAAATGGCCAAAAATCATCCTTTCCACTTGC  
 TAATTTTGTACAAATTCATCCTTACTAGTCTCCTGCCTGCTAGCAGAAATCTTTCCAGTCCT  
 CTTGTCTCCTTTGTTTGGCATCAGCAAGGGCTATGCTGTGATTCTGTCTCTGAGTGACTTG  
 GACCACTGACCCTCAGATTTCCAGCCTTAAAATCCACCTTCCTTCTCATGCGCCTCTCCGAA  
 TCACACCTGACTCTTCCAGCCTCCATGTCCAGACCTAGTCAGCCTCTCTCACTCCTGCCCC  
 TACTATCTATCATGGAATAACATCCAAGAAAGACACCTTGCATATTCTTTAGTTTCTGTTT  
 TGTCTCCACATATTCTCTTCAATGCTCAGGAAGCCTGCCCTGTGCTTGAGAGTTCAGGGC  
 CGGACACAGGCTCACAGGTCTCCACATTGGGTCCCTGTCTCTGGTGGCCACAGTGAGCTCCT  
 TCTTGGCTGAGCAGGCATGGAGACTGTAGTTTCCAGATTTCTGAAAAATAAAAGTTTACA  
 GCGTTATCTCTCCCAACCTCACTAA

**FIGURE 160**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71169

><subunit 1 of 1, 523 aa, 1 stop

><MW: 59581, pI: 8.68, NX(S/T): 1

MAGQRVLLLVGFLLPGVLLSEAAKILTISTVGGSHYLLMDRVSQILQDHGHNVTMLNHKRGPF  
FMPDFKKEEKSYQVISWLAPEDHQREFKKSFDFFLEETLGGRGKFENLLNVLEYLALQCSHF  
LNRKDIMDSLKNENFDMVIVETFDYCPFLIAEKLKGPVAILSTSFGSLEFGLPIPLSYVPV  
FRSLLTDHMDFWGRVKNFLMFFSFCRRQQHMQSTFDNTIKEHFTGSRPVLSHLLLKAELWF  
INSDFAFDFARPLLNTVYVGGLMEKPIKPVPQDLENFIAKFGDSGFVLVTLGSMVNTCQNP  
EIFKEMNNAFAHLPQGVIWKCQCShwPKDVHLAANVKIVDWLPQSDLLAHPSIRLFVTHGGQ  
NSIMEAIQHGVPVMVGIPLEFGDQPENMVRVEAKKFGVSIQLKKLKAETLALKMKQIMEDKRYK  
SAAVAASVILRSHPLSPTQRLVGWIDHVLQTGGATHLKPYVFQQPWHEQYLFDFVFLLGLT  
LGTLWLCGKLLGMAVWWLRGARKVKET

**Signal peptide:**

amino acids 1-19

**Transmembrane domain:**

amino acids 483-504

**FIGURE 161**

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GGGCTGTTGATTTGTGGGGGATTTTGAAGAGAGGAGGAATAGGAGGAAGGGGTTGAGGGGCT  
 GCCTCTGGCATATGCACACACTCACACATTCTGTACACCCCGTCACACACACATACCATGTT  
 CTCCATCCCCCAGGTCCAGCCCTCAGTGCTGTCCCATCCAGCAGGGCTACCCTGAAGCTCT  
 GGCTGCAGCCCTCCCGTCCAGTGGGCAGGCGGCTTCATCCCTCCTTTCTCTCCCAAAGCCCA  
 ACTGCTGTCACTGCATGCTCTGCCAAGGAGGAGGGAACTGCAGTGACAGCAGGAGTAAGAGT  
 GGGAGGCAGGACAGAGCTGGGACACAGGTATGGAGAGGGGGTTTCCAGCGAGCCTAGAGAGGGC  
 AGACTATCAGGGTGCCGGCGGTGAGAATCCAGGGAGAGGAGCGGAAACAGAAGAGGGGGCAGA  
 AGACCGGGGCACTTGTGGGTTGCAGAGCCCCCTCAGCC**ATG**TTGGGAGCCAAGCCCACTGGC  
 TACCAGGTCCCCTACACAGTCCCGGGCTGCCCTTGGTTCTGGTGCTTCTGGCCCTGGGGGCC  
 GGGTGGGCCCAGGAGGGGTCCAGAGCCCGTCTGCTGGAGGGGGAGTGCTGGTGGTCTGTGA  
 GCCTGGCCGAGCTGCTGCAGGGGGGCCCGGGGGAGCAGCCCTGGGAGAGGCACCCCCCTGGGC  
 GAGTGGCATTTGCTGCGGTCCGAAGCCACCACCATGAGCCAGCAGGGGAAACCGGCAATGGC  
 ACCAGTGGGGCCATCTACTTCGACCAGGTCTGGTGAACGAGGGCGGTGGCTTTGACCGGGC  
 CTCTGGCTCCTTCGTAGCCCCTGTCCGGGGTGTCTACAGCTTCCGGTTCCATGTGGTGAAGG  
 TGTACAACCGCCAAACTGTCCAGGTGAGCCTGATGCTGAACACGTGGCCTGTCTCTCAGCC  
 TTTGCCAATGATCCTGACGTGACCCGGGAGGCAGCCACCAGCTCTGTGCTACTGCCCTTGGA  
 CCCTGGGGACCGAGTGTCTCTGCGCCTGCGTCGGGGGAATCTACTGGGTGGTTGGAAATACT  
 CAAGTTTCTCTGGCTTCCCTCATCTTCCCTCTCT**TGAG**AGGCCCAAGTCTTTCAAGCACAGAAT  
 CCAGCCCCTGACAACTTTCTTCTGCCCTCTCTTGCCCCAGAAACAGCAGAGGCAGGAGAGAG  
 ACTCCCTCTGGCTCCTATCCCACCTCTTTGCATGGGACCCTGTGCCAAACACCCAAGTTTAA  
 GAGAAGAGTAGAGCTGTGGCATCTCCAGACCAGGCCCTTCCACCCACCCACCCCAAGTTACC  
 CTCCCAGCCACCTGCTGCATCTGTTCCCTGCCTGCAGCCCTAGGATCAGGGCAAGGTTTGGA  
 AGAAGGAAGATCTGCACTACTTTGCGGCCTCTGCTCCTCCGGTTCCCCACCCAGCTTCCT  
 GCTCAATGCTGATCAGGGACAGGTGGCGCAGGTGAGCCTGACAGGCCCCCACAGGAGCCAG  
 ATGGACAAGCCTCAGCGTACCCTGCAGGCTTCTTCCCTGTGAGGAAAGCCAGCATCACGGATC  
 TCAGCCAGCACCGTCAGAAGCTGAGCCAGCACCGTATGGGCTAGGGTGGGAGGCTCAGCCAC  
 AGGCAGAAGGGTGGGAAGGGCCTGGAGTCTGTGGCTGGTGAGGAAGGAAGGAGGGTGTATTG  
 TCTAGACTGAACATGGTACACATTCTGCATGTATAGCAGAGCAGCCAGCAGGTAGCAATCCT  
 GGCTGTCCTTCTATGCTGGATCCCAGATGGACTCTGGCCCTTACCTCCCCACCTGAGATTAG  
 GGTGAGTGTGTTTGGCTCTGGCTGAGAGCAGAGCTGAGAGCAGGTATACAGAGCTGGAAGTGG  
 ACCATGGAAAACATCGATAACCATGCATCCTCTTGCTTGGCCACCTCCTGAAACTGCTCCAC  
 CTTTGAAGTTTGAACTTTAGTCCCTCCACACTCTGACTGCTGCCTCCTTCCCTCCAGCTCTC  
 TCACTGAGTTATCTTCACTGTACCTGTTCCAGCATATCCCCACTATCTCTCTTTCTCCTGAT  
 CTGTGCTGTCTTATTCTCCTCCTTAGGCTTCCCTATTACCTGGGATTCCATGATTTCATTCTT  
 CAGACCCTCTCCTGCCAGTATGCTAAACCCTCCCTCTCTCTTTCTTATCCCGCTGTCCATT  
 GGCCAGCCTGGATGAATCTATCAATAAAACAAGTAGAGAATGGTGGTCAGTGAGACACTAT  
 AGAATTACTAAGGAGAAGATGCCTCTGGAGTTTGGATCGGGTGTACAGGTACAAGTAGGTA  
 TGTTGCAGAGGAAAATAAATATCAAACGTATACTAAATTAATAA

**FIGURE 162**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71180
><subunit 1 of 1, 205 aa, 1 stop
><MW: 21521, pI: 7.07, NX(S/T): 1
MLGAKPHWLPGPLHSPGLPLVLVLLALGAGWAQEGSEPVLLERGECLVVCEPGRAAAGGPGGA
ALGEAPPGRVAFAAVRSHHHEPAGETGNGTSGAIYFDQVLVNEGGGFDRASGSFVAPVRGVY
SFRFHVVVKVYNRQTVQVSLMLNTWPVISAFANDPDVTREAATSSVLLPLDPGDRVSLRLRRG
NLLGGWKYSSFSGLIFPL
```

**Signal peptide:**

amino acids 1-32

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



**FIGURE 163**

[illegible]

**FIGURE 164**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71184

><subunit 1 of 1, 388 aa, 1 stop

><MW: 43831, pI: 9.64, NX(S/T): 3

MKTLIAAYSGVLRGERQAEADRSQRSHGGPALSREGSGRWGTGSSILSALQDLFSVTWLNRS  
KVEKQLQVISVLQWVLSFLVLGVACSAILMYIFCTDCWLIADVLYFTWLVDWNTPKKGGRRS  
QWVRNWAVWRYFRDYFPIQLVKTHNLLTTRNYIFGYHHPHGIMGLGAFCNFSTEATEVSKKFP  
GIRPYLATLAGNFRMPVLRREYLMSSGICPVSRDTIDYLLSKNGSGNAIIIVVGGAESLSSM  
PGKNAVTLRNRKGFVKLALRHGADLVPIYSFGENEVYKQVIFEEGSWGRWVQKKFQKYIGFA  
PCIFHGRGLFSSDTWGLVPYSKPITTVVGEPITIPKLEHPTQQDIDLYHTMYMEALVKLFDK  
HKTKFGLPETEVLEVN

**Important features of the protein:**

**Transmembrane domain:**

amino acids 76-97

**N-glycosylation sites.**

amino acids 60-63, 173-176, 228-231

**N-myristoylation sites.**

amino acids 10-15, 41-46, 84-89, 120-125, 169-174, 229-234, 240-245, 318-323, 378-383

**FIGURE 165**

GGGCGGCGGGATGGGGGCCGGGGCGGGCGGGCGCCGCACTCGCTGAGGCCCCGACGCAGGGCCGGGCGGGCCCCA  
GGGCCGAGGAGCGCGGGCGGCCAGAGCGGGGCCGGAGGCGACGCCGGGGACGCCCGCGGACGAGCAGGTGGCG  
GCGGCTGCAGGCTTGTCCAGCCGGAAGCCCTGAGGGCAGCTGTTCCCACTGGCTCTGCTGACCTTGTGCCTTGGA  
CGGCTGTCTCAGCGAGGGGCCGTGCACCCGCTCCTGAGCAGCGCCATGGGCCTGCTGGCCTTCCTGAAGACCCA  
GTTTCGTGCTGCACCTGCTGGTTCGGCTTGTCTTCGTGGTGAGTGGTCTGGTCATCAACTTCGTCCAGCTGTGCAC  
GCTGGCGCTCTGGCCGGTCAGCAAGCAGCTCTACCGCCGCTCAACTGCCGCTCGCCTACTCACTCTGGAGCCA  
ACTGGTCATGCTGCTGGAGTGGTGGTCTGCACGGAGTGACACTGTTACGGACCAGGCCACGGTAGAGCGCTT  
TGGGAAGGAGCACGCAGTCATCATCTCAACCACAACTTCGAGATCGACTTCCTCTGTGGGTGGACCATGTGTGA  
GCGCTTCGGAGTGTGGGGAGCTCCAAGGTCCTCGCTAAGAAGGAGCTGCTCTACGTGCCCTCATCGGCTGGAC  
GTGGTACTTTCTGGAGATTGTGTTCTGCAAGCGGAAGTGGGAGGAGGACCGGGACACCGTGGTCAAGGGCTGAG  
GCGCCTGTGCGACTACCCGAGTACATGTGGTTTCTCCTGTACTGCGAGGGGACGCGCTTCACGGAGACCAAGCA  
CCGCGTTAGCATGGAGGTGGCGGCTGCTAAGGGGCTTCCTGTCTCAAGTACCACCTGCTGCCGCGGACCAAGGG  
CTTCACCACCGCAGTCAAGTGCCTCCGGGGGACAGTCGCAGCTGTCTATGATGTAACCCTGAACTTCAGAGGAAA  
CAAGAACCCGTCCCTGCTGGGGATCCTCTACGGGAAGAAGTACGAGGCGGACATGTGCGTGAGGAGATTTCCTCT  
GGAAGACATCCCGCTGGATGAAAAGGAAGCAGCTCAGTGGCTTCATAAACTGTACCAGGAGAAGGACGCGCTCCA  
GGAGATATATAATCAGAAGGGCATGTTTCAGGGGAGCAGTTTAAGCCTGCCCGGAGGCCGTGGACCCCTCCTGAA  
CTTCCTGTCTGGGCCACCATTCCTCTGTCTCCCTCTTCAGTTTTGTCTTGGGCGTCTTTGCCAGCGGATCACC  
TCTCCTGATCCTGACTTTCTTGGGGTTTGTGGGAGCAGCTTCCTTTGGAGTTCGAGACTGATAGGAGAATCGCT  
TGAACCTGGGAGGTGGAGATTGCAGTGAGCTGAGATGGCATCACTGTACTCCAGCCTAGGCAACAGAGCAAGACT  
CAGTCTCAAAAAAAAAAAAAAAAAACAAAAACCCAGAAATTCTGGAGTTGAACTGTGTAGTTACTGACATGAAAA  
ATTCACTAGAGGCTGAACAGCAGATTTGAGCAGGCAGAAAAAATCAGCAAGCTTGAAGATGGTACCTTGAGATT  
TTTCAGGCTAATGAAAAAGAATGAAGGAAAATTAACAGCCTCAGAGACCCATGGTGCACCGTCACACAAATCAA  
CATATGCATGATGAGAGTCCAGAGGAGAGAGAGAAAGGGTCAGAAAGAATGGCCACAAGCTGATGAAAAACA  
GTAACCTACCCACTCAGGAAGCTCAGTGAACCTCAATGAGGATGAATATCAGAGATCCACACCTAGATATTTTAT  
AATCAAAGTGTCAAATGACAAAGAATCTTGAAAGCAGCAAGAGATGAGCAACTTATCTTGTTCAAAGGATCTTG  
ATCAGATTAACAGCTCATTTCTCCTCAGAAATCATGGGAGCCAGGAGATAGTGGGATGAACACTGTTGAAGGCAA  
AACCTTCAACTGTAATTATTGGACTTTTGAGTCTTAGATGGTCTGACCTCTTGTCTTCAGGGACAGTTTTTCA  
ATTTAATCCCTAATAACAATTAGTCAAGCTTCCTTGACCTGTAGGAAGGCCTGTCTTTAGGCCGGGCACAGTGGC  
TTACACCTGTAATCCAGCACTTTGGGAGGCCAGACGGGTGGATCATTTGGGCTCAGGCTGATCTCAAACCTCT  
GAGTTCAGGTGATCTGCCCCCTCAGCCTCCCAAAGTGTTGTGATTGCAGGCGTGAGCCACTGCGCCTGGCCGGA  
ATTTCTTTTTAAGGCTGAATGATGGGGCCAGGCACGATGGCTCACGCCTGTGATCCCAAGTAGCTTGATTGTA  
AACATGCACCACCATGCCTGGCTAATTTTTGTATTTTTAGTAGAGACGTGTAGCCAGGCTGGTCTCGATCTCCT  
GACCTCAAGTGACCACCTGCCTCAGCCTCCCAAAGTACTGGGATTACAGGCGTGAGCCACTGTGCCTGGCCTTGA  
GCATCTTGTGATGTGCTTATTGGCCATTTGTATATCTTCTATCTTCTTTGGGGAAATGTCTGTTCAAGTCCTTG  
CCTTTTTTAAATTTTTATTATTTATTTATTTATTTATTTTGTAGACAGGGTCTTGTCTGTGTTGCCAGGCTGGAGTA  
CAGTGGCACAGTCTTGGCTCACTGCAGCCTCGACCTCCTGGGCTGCAGTGATCCTCCACCTCAGCCTCCCTTGT  
AGCTGTATTTTTTTGTATTTTGTATTTTGTAGCTGTAGTTTTTGTATTTTTTGTGGAGACAGCATTTACCATGA  
TGCCAGGCTGGTCTTGAACCTCTGAGCTCAAGTGATCTGCCTGCTTCAGCCTCCCAAAGTGCTGGGATTACAGA  
CATGAGCCACTGCACCTGGCAAACCTCCCAAATTAACACACACACACAAAAAACCTGATTCAAAATGGGCA  
GAGGGGCCGGGTGTGGCCCCAACTACCAGGGAGACTGAAGTGGGAGGATCGCTTGGGCATGAGAAGTCAGGCTG  
CAGTGAGTCGAGGTTGTGCGACTGCATTCCAGCCTGGACAACAGAGTGAGACCCTGTCTC

**FIGURE 166**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71213

><subunit 1 of 1, 368 aa, 1 stop

><MW: 42550, pI: 9.11, NX(S/T): 1

MGLLAFLKTQFVLHLLVGFVVFVVSGLVINQVQLCTLALWPVSKQLYRRRLNCRLAYSLWSQLV  
MLLEWWSCTECTLFTDQATVERFGKEHAVIILNHNFEIDFLCGWTMCERFGVLGSSKVLAKK  
ELLYVPLIGWTWYFLEIVFCKRKWEEDRDTVVEGLRRLSDYPEYMWFLLYCEGTRFTETKHR  
VSMEVAAAKGLPVLKYHLLPRTKGFTTAVKCLRGTVAAVYDVTLNFRGNKNPSLLGILYGKK  
YEADMCVRRFPLEDIPLDEKEAAQWLHKLYQEKDALQEIYNQKGMFPGEQFKPARRPWTLN  
FLSWATILLSPLFSFVLGVFASGSPLLILTFLGFVGAASFGVRRRLIGESLEPGRWRLQ

**Important features of the protein:**

**Signal peptide:**

amino acids 1-25

**Transmembrane domains:**

amino acids 307-323, 335-352

**Tyrosine kinase phosphorylation sites.**

amino acids 160-168, 161-169

**FIGURE 167**

GATATTCTTTATTTTAAAGAATCTGAAGTACTATGCATCACTCCCTCCAATGTCCTGGGGCA  
GCCACCAGGCATATTCATCTTTGTGTGTGTTTTCTTTGCTTTAGCACTGGGGCACTTCTT  
GCTTATTTCTTTGGTAGGAAAGGGGCTCAGTTTGTCTTGTGGGGTTGGTGGCAGGCAGGCCG  
GCTTACGCCTGATACGGCCCTGGGTTAGAAGGGAAGGGAAGATAAACTTTTATACAAATGGG  
GATAGCTGGGGTCTGAGACCTGCTTCCTCAGTAAAATTCCTGGGATCTGCCTATACCTTCTT  
TTCTCTAACCTGGCATAACCTGCTTAAAGCCTCTCAGGGCTTCTCTCTGTTCTTAGGATCAA  
AGTATTTAGAGCTACAAGAGCCCTCATGGTCTGGCCCCCTGCCCCCTGGCCAGCTTCATTGT  
ACATGTGGTGTTCTCTTGTGCTTCCTGTAATGTGGTATGCCATGGGGTCTTTGCACAAGCCT  
TTCCTCTTTGGCTGGACACTGTTCCCTGCCCCCCCCATACTCTTCCTACTTAATATGTAGTC  
ATCCTGCAGATTTCAATTCTAACATCATTTTCTCCAGGGATCCTGGCCTGACAGAATCTCAT  
CTTGTTTAATGCTCTCATAAGACCACTTGTTTCCCTTTTGCAGCACTTGCCACTCAGTTGTA  
TCTTTATGTGCGTTTGTGGTTGTATGGGTTGTGTCTGTTCCCCAGAATGCCAGCTCTGAGC  
TGCGTGAGGGTCAAGGGCATTGCTGTGCCTGCCAGGTATAGTGCCTACATGTGGTGGGTGCT  
CATGTTTTAGAGACTAAATGGAGGAGGAGATGAGGAAAAGATTGAAATCTCTCAGTTCACCA  
GATGGTGTAGGGCCCAGCATTGTAAATTCACACGTTGACTGTGCTTGTGAATTATCTGGGGA  
TGCAGGTCCTGATTCACTAGGCCAGGTTGGGCATCTCTAACAACTCCCACGTGATGCTGA  
TGCTGGTCCTATGAACTATACTAAATAGTAAGAATCTATGGAGCCAGGCTGGGCATGGTSGC  
TCACACCTATGATCCCAGCACTTTGGGAGGCTGAGGCAGGCTGATCACCTGGAGTCAGGATT  
TCAAGACTAGCCTGGCCAACATGGTGGAACCCCATCTGTACTAAAAATACACAAATTAGCTG  
GGCATGGTGGCACATGCCTGTAGTCCCAGCTACTTGGGAGGCTGAAGCAAGAGAATCGCTTG  
AACCTGGGAGGCGGAGGTTGCAGTGAGCCGAGATCAGGCCACTGTATTCCAACCAGGGTGAC  
AGAGTGAGACTCTATGTCCAAAAAAAAAAAAA

**FIGURE 168**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71234

><subunit 1 of 1, 143 aa, 1 stop

><MW: 15624, pI: 9.58, NX(S/T): 0

MHSLQCPGAATRIHLVCFSFALALGHFLLISLVGKGLSLSCGVGGRQAGLRLIRPWVRR  
EGKINFYTNGDSWGLRPASSVKFLGSAYTFFSLTWHTLLKASQGFSLFLGSKYLELQEPSWS  
GPCPPGQLHCTCGVLLSFL

**Important features of the protein:****Signal peptide:**

amino acids 1-28

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250

**FIGURE 169**

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGA  
 TTATAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTCAGCAACTAAAAAAGCCAC  
 AGGAGTTGAACTGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTAC  
 ATTAAATCTGTTTTTTTGTCTCTTGTAACTAGCCTTTACCTTCCTAACACAGAGGATCTGT  
 CACTGTGGCTCTGGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCAC  
 ACCGTCCCCTCGAAGCCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTC  
 ACCAACTGTCTCACGTCTGGAGGCACTGACTCGGGCAGTGACAGGTAGCTGAGCCTCTTGGA  
 GCTGCGGCTTTCAAGGTGGGCCTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTT  
 CATAGGCGATGGCTCCCCTGCCCAGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGG  
 CCAGGACGGGCCGTGGACACCTGCTCAGAAGCAGTGGGTGAGACATCACGCTGCCCCGCCAT  
 CTAACCTTTTCATGTCCTGCACATCACCTGATCCATGGGCTAATCTGAACTCTGTCCCAAGG  
 AACCAGAGCTTGAGTGAGCTGTGGCTCAGACCCAGAAGGGGTCTGCTTAGACCACCTGGTT  
 TATGTGACAGGACTTGCATTCTCCTGGAACATGAGGGAACGCCGGAGGAAAGCAAAGTGGCA  
 GGAAGGAACTTGTGCCAAATTATGGGTCAGAAAAGATGGAGGTGTTGGGTTATCACAAGGC  
 ATCGAGTCTCCTGCATTCACTGGACATGTGGGGGAAGGGCTGCCGATGGCGCATGACACACT  
 CGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGATCCACGTACCAGCTGCTG  
 AAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAATCTGCGATCACCAG  
 CCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCCCTCCTTCCCTC  
 TGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGGGGCTAATG  
 GCTCAGTGTTGGCCCAGGAGGTCAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCTGTGCG  
 AACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGGCT  
 CAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGT  
 TCAAATGATCTCCAAGGGCCCTTATACCCCAGGAGACTTTGATTTGAATTTGAAACCCCAA  
 TCCAAACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATG  
 CCAACATTTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTG  
 GCCAACATGGTGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGT  
 GTGTGCCTGTATCCCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGT  
 GAAGGAGGCTGAGACAGGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGA  
 AAAAATAAAAAAAGAATTATGGTTATTTGTAA





**FIGURE 171**

GCGGGCCCGCGAGTCCGAGACCTGTCCCAGGAGCTCCAGCTCACGTGACCTGTCACTGCCTC  
 CCGCCGCCTCCTGCCCCGCGCC**ATG**ACCCAGCCGGTGCCCCGGCTCTCCGTGCCCCGCCGCGCT  
 GGCCCTGGGCTCAGCCGCACTGGGCGCCGCCCTTCGCCACTGGCCTCTTCCTGGGGAGGCGGT  
 GCCCCCATGGCGAGGCCGGCGAGAGCAGTGCCTGCTTCCCCCGAGGACAGCCGCCTGTGG  
 CAGTATCTTCTGAGCCGCTCCATGCGGGAGCACCCGGCGCTGCGAAGCCTGAGGCTGCTGAC  
 CCTGGAGCAGCCGCAGGGGGATTCTATGATGACCTGCGAGCAGGCCCAGCTCTTGGCCAACC  
 TGGCGCGGCTCATCCAGGCCAAGAAGGCGCTGGACCTGGGCACCTTCACGGGCTACTCCGCC  
 CTGGCCCTGGCCCTGGCGCTGCCCCGCGGACGGGCGCGTGGTGACCTGCGAGGTGGACGCGCA  
 GCCCCCGGAGCTGGGACGGCCCCTGTGGAGGCAGGCCGAGGCGGAGCACAAGATCGACCTCC  
 GGCTGAAGCCCGCCTTGGAGACCCTGGACGAGCTGCTGGCGGCGGGCGAGGCCGGCACCTTC  
 GACGTGGCCGTGGTGGATGCGGACAAGGAGAACTGCTCCGCCTACTACGAGCGCTGCCTGCA  
 GCTGCTGCGACCCGGAGGCATCCTCGCCGTCCTCAGAGTCCTGTGGCGCGGGAAGGTGCTGC  
 AACCTCCGAAAGGGGACGTGGCGGCCGAGTGTGTGCGAAACCTAAACGAACGCATCCGGCGG  
 GACGTCAGGGTCTACATCAGCCTCCTGCCCCGATGGACTCACCTTGGCCTTCAAGAT  
**CTAG**GGCTGGCCCCTAGTGAGTGGGCTCGAGGGAGGGTTGCCTGGGAACCCAGGAATTGAC  
 CCTGAGTTTTAAATTCGAAAATAAAGTGGGGCTGGGACACAAAAAAAAAAAAAAAAAAAAA

T  
O  
T  
A  
L  
S  
E  
Q  
U  
E  
N  
C  
E  
S  
= 693100

**FIGURE 172**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71282

><subunit 1 of 1, 262 aa, 1 stop

><MW: 28809, pI: 8.80, NX(S/T): 1

MTQVPVRLSVPAALALGSAALGAAFATGLFLGRRCPPWRGRREQCLLPEDSRLWQYLLSRS  
MREHPALRSLRLLTLEQPQGDSMMTCEQAQLLANLARLIQAKKALDLGTFTGYSALALALAL  
PADGRVVTCEVDAQPPELGRPLWRQAEAEHKIDLRLKPALETLDELLAAGEAGTFDVAVVDA  
DKENCSAYYERCLQLLRPGGILAVLRVLWRGKVLQPPKGDVAAECVRNLNERIRRDVRVYIS  
LLPLGDGLTLAFKI

**Important features of the protein:****Signal peptide:**

amino acids 1-25

**Transmembrane domains:**

amino acids 8-30, 109-130

**N-glycosylation site.**

amino acids 190-193

**Tyrosine kinase phosphorylation site.**

amino acids 238-246

**N-myristoylation sites.**

amino acids 22-27, 28-33, 110-115, 205-210, 255-260

**Amidation sites.**

amino acids 31-34, 39-42

**FIGURE 173**

[illegible]

**FIGURE 174**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71286

><subunit 1 of 1, 671 aa, 1 stop

><MW: 74317, pI: 7.61, NX(S/T): 0

MPHAFKPGDLVFAKMKGYPHWPARIDDIADGAVKPPPNKYPIFFFGTHETAFLGPKDLFPYD  
KCKDKYGKPNKRKGFNEGLWEIQNNPHASYSAPPPVSSSDSEAPEANPADGSDADEDDEDRG  
VMAVTAVTATAASDRMESDSDSKSSDNSGLKRKTPALKMSVSKRARKASSDLDQASVSPSE  
EENSESSSESEKTSQDFTPEKKA AVRAPRRGPLGGRKKKKAPSASDSDSKADSDGAKPEPV  
AMARSASSSSSSSSSSSDSDSVKPKPRGRKPAEKPLPKPRGRKPKPERPPSSSSSDSDSDEV  
DRISEWKRRDEARRRELEARRRREQEEELRRLREQEKEEKERRRERADRGEAERGSGGSSGD  
ELREDDEPVKKRGRKGRGRGPPSSSDSEPEAELEAKKSAKKPQSSSTEPARKPGQKEKRV  
RPEEKQQAQKPVKVERTKRSEGFMSMDRKVEKKKEPSVEEKLQKLHSEIKFALKVDS PDVKRC  
LNALEELGTLQVTSQILQKNTDVVATLKKIRRYKANKDVMKAAEVYTRLKSRVLGPKIEAV  
QKVNKAGMEKEKAEKLAGEEELAGEEAPQEKAEDKPSTDLSAPVNGEATSQKGESAEDEKEHE  
EGRDSEEGPRCGSSEDLHDSVREGPDLDRPGSDRQERERARGDSEALDEES

**Signal peptide:**

amino acids 1-13

**FIGURE 175**

GTTGGTTCTCCTGGATCTTCACCTTACCAACTGCAGATCTTGGGACTCATCAGCCTCAATAATTATATTAAATTA  
 ACACCATTTGAAAGAGAACATTGTTTTTCATCATGAATGCTAATAAAGATGAAAGACTTAAAGCCAGAAGCCAAGA  
 TTTTCACCTTTTTCTGCTTTGATGATGCTAAGCATGACCATGTTGTTTCTTCCAGTCACTGGCACTTTGAAGCA  
 AAATATTTCCAAGACTCAAGCTAACCTACAAAGACTTGCTGCTTTCAAATAGCTGTATTCCTTTTTTGGGTTCATC  
 AGAAGGACTGGATTTTCAAACCTTCTCTTAGATGAGGAAAGAGGCAGGCTGCTCTTGGGAGCCAAAGACCACAT  
 CTTTCTACTCAGTCTGGTTGACTTAAACAAAAATTTTAAAGAAGATTTATTGGCCTGCTGCAAAGGAACGGGTGGA  
 ATTATGTAAATTAGCTGGGAAAGATGCCAATACAGAATGTGCAAATTTTCATCAGAGTACTTCAGCCCTATAACAA  
 AACTCACATATATGTGTGTGGAACCTGGAGCATTTTCATCCAATATGTGGGTATATTGATCTTGGAGTCTACAAGGA  
 GGATATTATATTCAAACCTAGACACACATAATTTGGAGTCTGGCAGACTGAAATGTCTTTTCGATCCTCAGCAGCC  
 TTTTGCTTCAGTAATGACAGATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCCCTTGGCAAAGATACTGCATT  
 CACTCGATCCCTTGGGCCCTACTCATGACCACCCTACATCAGAAGTACATTTTCAGAGCACTACTGGCTCAATGG  
 AGCAAAATTTATTGGAACCTTTCTTCATACCAGACACCTACAATCCAGATGATGATAAAATATATTTCTTCTTTTCG  
 TGAATCATCTCAAGAAGGCAGTACCTCCGATAAAACCATCCTTTCTCGAGTTGGAAGAGTTTGTAAAGAATGATGT  
 AGGAGGACAACGCAGCCTGATAAACAGTGGACGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTCCTGGAAG  
 TGATGGGGCAGATACTTACTTTGATGAGCTTCAAGATATTTATTTACTCCCACAAGAGATGAAAGAAATCCTGT  
 AGTATATGGAGTCTTTTACTACAACCAGCTCCATCTTCAAAGGCTCTGCTGTTTGTGTGTATAGCATGGCTGACAT  
 CAGAGCAGTTTTTAAATGGTCCATATGCTCATAAGGAAAGTGCAGACCATCGTTGGGTGCAGTATGATGGGAGAAT  
 TCCTTATCCACGGCCTGGTACATGTCCAAGCAAAACCTATGACCCACTGATTAAGTCCACCCGAGATTTTCCAGA  
 TGATGTTCATCAGTTTTCATAAAGCGGCAGTCTGTGATGTATAAGTCCGTATACCCAGTTGCAGGAGGACCAACGTT  
 CAAGAGAATCAATGTGGATTACAGACTGACACAGATAGTGGTGGATCATGTTCATTGCAGAAGATGGCCAGTACGA  
 TGTAATGTTTTCTTGGAACAGACATTTGGAACCTGTCTCAAAGTTGTGAGCATTTCAAAGGAAAAGTGAATATGGA  
 AGAGGTAGTGTCTGGAGGAGTTGCAGATATTTCAAGCACTCATCAATCATCTTGAACATGGAATTTGTCTCTGAAGCA  
 GCAACAATTGTACATTTGGTTCCTGAGATGGATTAGTTGAGCTCTCCTTGCACAGATGCGACACTTATGGGAAAGC  
 TTGCGCAGACTGTTGTCTTGGCAGAGACCCCTACTGTGCCTGGGATGGAATGCATGCTCTCGATATGCTCCTAC  
 TTCTAAAAGGAGAGCTAGACGCCAAGATGTAAAATATGGCGACCCAATCACCCAGTGTCTGGGACATCGAAGACAG  
 CATTAGTCATGAAACTGCTGATGAAAAGGTGATTTTTTGGCATTGAATTTAACTCAACCTTTCTGGAATGTATACC  
 TAAATCCCAACAAGCAACTATTAAATGGTATATCCAGAGGTGAGGGATGAGCATCGAGAGGAGTTGAAGCCCGA  
 TGAAAGAATCATCAAAACGGAATATGGGCTACTGATTGCAAGTTTGCAGAAGAAGGATTCTGGGATGTATTACTG  
 CAAAGCCCAGGAGCACACTTTTCATCCACACCATAGTGAAGCTGACTTTGAATGTTCATTGAGAATGAACAGATGGA  
 AAATACCCAGAGGGCAGAGCATGAGGAGGGGCGAGGTCAAGGATCTATTGGCTGAGTCACGGTTGAGATACAAAGA  
 CTACATCCAAATCCTTAGCAGCCCAAACTTCAGCCTCGACCAGTACTGCGAACAGATGTGGCACAGGGAGAAGCG  
 GAGACAGAGAACAAGGGGGGCCAAAGTGAAGCTCAGGAAATGAAGAAGAAACGAAATCGAAGACATCA  
 CAGAGACCTGGATGAGTCTCCTAGAGCTGTAGCCAGTGTATTTTCTACTTAATTTAAAGAAAAGAATTCCTTACC  
 TATAAAAACATTGCCTTCTGTTTTGTATATCCCTTATAGTAATTCATAAATGCTTCCCATGGAGTTTTTGCTAAGG  
 CACAAGACAATAATCTGAATAAGACAATATGTGATGAATATAAGAAAGGGCAAAAATTCATTTGAACCAGTTTT  
 CCAAGAACAATCTTGCACAAGCAAAGTATAAGAATTATCCTAAAAATAGGGGGTTTACAGTTGTAAATGTTTTA  
 TGTTTTGAATTTTGGAAATTTATTGTTCATGTAAATAGTTGAGCTAAGCAAGCCCCGAATTTGATAGTGTATAAGGT  
 GCTTTATTTCCCTCGAATGTCCATTAAGCATGGAATTTACCATGCAGTTGTGCTATGTTCTTATGAACAGATATAT  
 CATTCTATTGAGAACCAGCTACCTTGTGGTAGGGAATAAGAGGTCAGACACAAATTAAGACAACCTCCCATATC  
 AACAGGAACCTTTCTCAGTGAGCCATTCACTCCTGGAGAATGGTATAGGAATTTGGAGAGGTGCATTATTTCTTTC  
 TGGCCACTGGGGTTAAATTTAGTGTACTACAACATTGATTTACTGAAGGGCACTAATGTTTCCCCCAGGATTTCT  
 ATTGACTAGTCAGGAGTAACAGGTTACAGAGAGAAGTTGGTGTCTAGTTATGTGTTTTTTAGAGTATATACTAA  
 GCTCTACAGGGACAGAATGCTTAATAAATACTTTAATAAGATATAGGAAAATATTTTAAATAAAACAAGGAAAACA  
 TAATGATGTATAATGCATCCTGATGGGAAGGCATGCAGATGGGATTTGTTAGAAGACAGAAGGAAAGACAGCCAT  
 AAATTCTGGCTTTGGGGAAAACCTCATATCCCCATGAAAGGAAGAACAATCACAAATAAAGTGAGAGTAATGTAA  
 TGGAGCTCTTTTCACTAGGGTATAAGTAGCTGCCAATTTGTAATTCATCTGTTAAAAAAAATCTAGATTATAACA  
 AACTGCTAGCAAAATCTGAGGAAACATAAATCTTCTGAAGAATCATAGGAAGAGTAGACATTTTATTTATAACC  
 AATGATATTTTCAGTATATATTTTCTCTCTTTTAAAAAATATTTATCATACTCTGTATATTATTTCTTTTACTGC  
 CTTTATTTCTCCTGTATATTGGATTTTGTGATTATATTTGAGTGAATAGGAGAAAACAATATATAACACACAGA  
 GAATTAAGAAAATGACATTTCTGGGGAGTGGGGATATATATTTGTTGAATAACAGAACGAGTGTAAATTTTAAAC  
 AACGGAAAGGGTTAAATTAACCTTTGACATCTTCACTCAACCTTTTCTCATTGCTGAGTTAATCTGTTGTAAAT  
 GTAGTATTGTTTTTGTAAATTAACAATAAATAAGCCTGCTACATGT

**FIGURE 176**

```

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71883
><subunit 1 of 1, 777 aa, 1 stop
><MW: 89651, pI: 7.97, NX(S/T): 3
MNANKDERLKARSQDFHLFPALMMLSMTMLFLPVTGTLKQNI PRLKLT YKDLLLSNSCIPFL
GSSEGLDFQTLLLDEERGRLLLGA KDHI FLLSLVDLNKNFKKIYWPA AKERVELCKLAGKDA
NTECANFIRVLQPYNKTHIYVCGTGAFHPICGYIDLGVYKEDIIFKLDTHNLESGRLKCPFD
PQQPFASVMTDEYLYSGTASDFLGKDTA FTRSLGPTH DHYIRTDISEHYWLN GAKFIGTFF
IPDTYNPDDDKIYFF FRESSQEGSTSDKTILSRVGRVCKNDVGGQ RSLINKWTTFLKARLIC
SIPGSDGADTYFDELQDIYLLPTRDERNPVVYGVFTTTSSIFKGSAVCVYSMADIRAVFNGP
YAHKESADHRWVQYDGRIPYPRPGT CPSKTYDPLIKSTRDFPDDVISFIKRHSVMYKSVYPV
AGGP TFKRINVDYRLTQIVVDHVIAEDGQYDVMFLGTDIGTVLKVVSISKEKWNMEEVVLEE
LQIFKHSSIILNMELSLKQQQLYIGSRDGLVQLSLHRCDTYGKACADCCCLARDPYCAWDGNA
CSRYAPTSKRRARRQDVKYGDPITQCWDIEDSISHETADEKVI FGIEFNSTFLECI PKSQQA
TIKWYIQRSGDEHREELKPDERIIKTEYGLLIRSLQKKDSGMY YCKAQEHTFIHTIVKLTLN
VIENEQMENTQRAEHEEGQVKDLLAESRLRYKDYIQILSSPNFSLDQYCEQMWHREKRRQRN
KGGPKWKHMQEMKKKRNRRHHRDLDELPRAVAT

```

**Important features of the protein:****Signal peptide:**

amino acids 1-36

**N-glycosylation sites.**

amino acids 139-142, 607-610, 724-727

**Tyrosine kinase phosphorylation site.**

amino acids 571-576

**Gram-positive cocci surface proteins 'anchoring' hexapeptide.**

amino acids 32-37

**FIGURE 177**

CCCTGACCTCCCTGAGCCACACTGAGCTGGAAGCCGAGAGGTCATCCTGGAGCATGCCACCCGCGGGGAGCAGA  
CAACCTCCCAGGTAAGCTGGGAGCAAGACCTGAAGCTGTTTTCTTCAGGAGCCTGGTGTATTTTTCCCCACCCAC  
CTCAGCAGTTTTAGCCAGCAGGGACTGATCAGGTGTGTCTCTGGAGTGGGGAGCAGAAAGCGTGGCTGGCAAGA  
GTGGCCTGGAGAAAGAGGTTAGCGCTTGACCAGCCGAGCTGCCCGTGAATAAGATCCAGAACCATTGGGCATC  
GGGTGAGGTGGGGGGGCACAGGTGTATGTGCACCTTCTTGTCTCAGCAAGAAGAGCTGAGAGAGGGGATCTTGG  
AGCCATTGAGGTGTATGGAGCTACAGAGGGGAGGGAAAGGTATTTTAAGGTAACAGTGTGGCACAATAGTTAA  
GAGCACAGTTTTTTGGAGCTAGACCGACATAGGTTCAAATTTCTTCTGTGTGCTTCTTAGTTCTGTAGCCCCAGGT  
AAGGGAGTGACTTAACCTCTCTGGACTTCAATTTCTCATCTAAAGTAGGGCCAATAATAGCACCCACCTCAT  
AGGGAAGATTAAATGACATAATGTATGTGATGCACTAGCAAAAGTACCAGTCCCATAGTAAGTCATGCCCCACAG  
TATTTCCACCCACCCCTGTTCTCTGCCTTCCCAACCAGGTACTGCAACGACTGGAGCAGAGGCGGCAGCAGGCTT  
CAGAGCGGGAGGCTCCAAGCATAGAACAGAGGTTACAGGAAGTGCAGAGAGCATCCGCCGGGCACAGGTGAGCC  
AGGTGAAGGGGGCTGCCCGGCTGGCCCTGCTGCAGGGGGCTGGCTTAGATGTGGAGCGCTGGCTGAAGCCAGCCA  
TGACCCAGGCCAGGATGAGGTGGAGCAGGAGCGGCGGCTCAGTGAGGCTCGGCTGTCCAGAGGGACCTCTCTC  
CAACCGCTGAGGATGCTGAGCTTTCTGACTTTGAGGAATGTGAGGAGACGGGAGAGCTCTTTGAGGAGCCTGCCC  
CCCAAGCCCTGGCCACGAGGGCCCTCCCTGCCCTGCACACGTGGTATTTTCGCTATCAGGCAGGGCGTGAGGATG  
AGCTGACAATCACGGAGGGTGAGTGGCTGGAGGTATAGAGGAGGGAGATGCTGACGAATGGGTCAAGGCTCGGA  
ACCAGCACGGCGAGGTAGGCTTTGTCCCTGAGCGATATCTCAACTTCCCGGACCTCTCCCTCCAGAGAGCAGCC  
AAGACAGTGACAATCCCTGCGGGGCAGAGCCACAGCATTCTGGCACAGGCCCTGTACAGCTACACCGGACAGA  
GTGCAGAGGAGCTGAGCTTCCCTGAGGGGGCACTCATCCGTCTGCTGCCCCGGGCCCAAGATGGAGTAGATGACG  
GCTTCTGGAGGGGAGAATTTGGGGGCGGTGTTGGGGTCTTCCCTCCCTGCTGGTGGAGAGCTGCTTGGCCCCC  
CAGGGCCACCTGAACTCTCTGACCCCTGAACAGATGCTGCCCTTCCCTTCTCCTCCAGCTTCTCCCCACCTGCAC  
CTACCTCTGTGTGGATGGCCCCCTGCACCTGTCTGCTGGGGACAAAGCCCTGGACTTCCCTGGGTTCTCTGG  
ACATGATGGCACCTCGACTCAGGCCGATGCGTCCACCACCTCCCCCGCCGGCTAAAGCCCCGGATCCTGGCCACC  
CAGATCCCCCTCACCCTGAAGGCCAGGGAAGCCTTGACCCCCAGTGATGCTGCTGTCCCTATCTTCAAGCTGTGAGA  
CCACACCATCAATGATCCAGAGCAACACAGCCAAAAGCTGGAATCGCCCTTATTTCCACCCTCACCTCCAAGGGT  
GGAAACTTGCCCCCTCCCATTTCTAGAGCTGGAACCCACTCCTTTTTTTCCATTGTTCTATCATCTCTAGGACC  
GGAACTACTACCTTCTCTTCTGTCTAGACTGGAACCCACTCCTTTTTTTCCATTGTTCTATCATCTCTAGGACC  
ATCCATCAAGGTCTCTAGTAGTTCTGGCCACCTCTTTCCCCACCTGGCTCCATGACCCACCCCACTCTGGATG  
CCAGGTCACTGGGGTGGGCTGGGGAGAGGAACAGGCCCTTGGGAATCAGGAGCTGGAGCCAGGATGCGAAGCAG  
CTGTAATGGTCTGAGCGGATTTATTGACAATGAATAAAGGGCACGAAGGCCAGGCCAGGGCCTGGGCCTCTTGTG  
CTAAGAGGGCAGGGGGCCTACGGTGCTATTGCTTTAGGGGCCACACAGGGCAGGGGCTGCTCCAGCTGCCAC  
GCTCTATCATATGGAGCGAGGTGTTGGGGAGCGAGGCTGTTGCAGGCAGGGGAAGGAGAAGAGAC  
TGAGGGGCTGTGACCTCTCTGAGGCCCCAGCCTGAGACTGTGCAACTCCAGGTGGAAGTAGAGCTGGTCCCTC  
AGCTGGGGGCACTGCTGTCCAGTGGAGGGGAGGGCTTTACGCCCCACCCACCCCTGGCCCTGCCAGCTGGTAG  
TCCATCAGCACAATGAAGGAGACTTGGAGAAGAGGAAGAATAACACTGTTGCTTCTGTTCAAGCTGTGTCCAGC  
TTTTCCCTGGGGCTCCAGGACCTTCCCTACCTCCACCACCAACCAAGGATTTATAGCAAAGGCTAAGCCTGC  
AGTTTACTCTGGGGGTTGAGGGAGCCGAAAGGCTTAAATAGTTTAAAGTAGGTGATGGGAAGATGAGATTACCTCA  
TTTAGGGCTCAGGCAGACTCACCTCACATACTCCCTGCTCCCTGTGGTAGAGACACCTGAGAGAAAGGGAGGGG  
TCAACAATGAGAGACCAGGAGTAGGTCTATCAGTGCCCCCAGAGTAGAGAGCAATAAGAGCCCAGCCAGTGC  
AGTCCCGCTGTGTTTTCTACCTGGTGATCAGAAGTGTCTGGTTTGGCTGGCTGCCATTTGCCTCTTGAGTGG  
GCAGCCCTGGGCTTGGGGCCCTCCCTCCGGCCCTCAGTGTTGGCTCTGCAGAAGCTCTGGGGTTCCCTTCAAGTG  
CAGGAGGGGTAGGCTGCTGTCCCTGAGTCTCCATTCTGTACTGGGGGGCTGGCTAGGACCTGGGGCTGTGGCC  
TCTCAGGGGGCAGCCTCTCCATGGCAGGCATCCCTGCCTTGGGCTGCCCTCCCCAGACCCCTGACCACCCCTG  
GGTCTGTCCCCCACCAGAGCCCCAGCTCCTGTCTGTGGGGGAGCCATCACGGTGTTCGTGCAGTCCATAGCGCT  
TCTCAATGTGTGTACCCGGAACCTGGGAGGGGAGGGAACACTGGGGTTTAGGACCACAACCTCAGAGGCTGCTTG  
GCCCTCCCCCTCTGACCAGGGACATCCTGAGTTTTGGTGGCTACTTCCCTCTGGCCTAAGGTAGGGGAGGCTTCTC  
AGATTGTGGGGCACATTGTGTAGCCTGACTTCTGCTGGAGCTCCAGTCCAGGAGGAAAGAGCCAAGGCCACTT  
TTGGGATCAGGTGCCTGATCACTGGGCCCCCTACCTCAGCCCCCTTTCCCTGGAGCACCTGCCCCACCTGCCCCA  
CAGAGAACACAGTGGTCTCCCTGTCCGGGGGCGGCTTTTTCCTTCCCTGGAGCGTCCCTGACGGACAAGTGGAG  
GCCTCTTGTGCGGCTGCAATGGATGCAAGGGGCTGCAGAGCCAGGTGCACTGTGTGATGATGGGAGGGGGCTC  
CGTCTGTCAGGCTGGAGGTGGCATCCACACTGGACAGCAGGAGGAGGGGAGTGAGGGTAACATTTCCATTTCCCT  
TCATGTTTTGTTTTCTTACGTTCTTTAGCATGCTCCTTAAAACCCAGAACCCCAATTTCCCCAAGCCCCATTT  
TTTCTTGTCTTTATCTAATAAATCAATATTAAG

**FIGURE 178**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73401

><subunit 1 of 1, 370 aa, 1 stop

><MW: 40685, pI: 4.53, NX(S/T): 0

MQLAKYQSHSKSCPTVFPPTPVLCLPNQVLQRLEQRRQQASEREAPSIEQRLQEVRESIRRA  
QVSQVKGAAARLALLQGAGLDVERWLKPAMTQAQDEVEQERRLSEARLSQRDLSPTAEDAELS  
DFEECEETGELFEPPAPQALATRALPCPAHVVFYQAGREDELTITEGEWLEVIEEGDADEW  
VKARNQHGEVGVFPERYLNFPDLSLPESQSDSNPCGAEPATAFLAQALYSYTGQSAEELSFP  
EGALIRLLPRAQDGVDDGFWRGEFGGRVGVFPSLLVEELLGPPGPPPELSDPEQMLPSPSPPS  
FSPPAPTSVLDGPPAPVLPGDKALDFPGFLDMMAPRLRPMRPPPPPPAKAPDPGHPDPLT

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



**FIGURE 179A**

CACAGGGAGACCCACAGACACATATGCACGAGAGAGACAGAGGAGGAAAGAGACAGAGACAAAGGCACAGCGGAA  
GAAGGCAGAGACAGGGCAGGCACAGAAGCGGCCAGACAGAGTCTACAGAGGGAGAGGCCAGAGAAGCTGCAGA  
AGACACAGGCAGGGAGAGACAAAGATCCAGGAAAGGAGGGCTCAGGAGGAGAGTTTGGAGAAGCCAGACCCCTGG  
GCACCTCTCCCAAGCCCAAGGACTAAGTTTTCTCCATTTCTTTAACGGTCTCAGCCCTTCTGAAAACCTTTGCC  
TCTGACCTTGGCAGGAGTCCAAGCCCCCAGGCTACAGAGAGGAGCTTTCCAAAGCTAGGGTGTGGAGGACTTGGT  
GCCCTAGACGGCCTCAGTCCCTCCCAGCTGCAGTACCAGTGGC**ATG**TCCAGACAGGCTCGCATCCCCGGGAGGGG  
CTTGGCAGGGCGCTGGCTGTGGGGAGCCCAACCCTGCCTCCTGCTCCCCATTGTGCCGCTCTCCTGGCTGGTGTG  
GCTGCTTCTGCTACTGCTGGCCTCTCTCCTGCCCTCAGCCCGGCTGGCCAGCCCCCTCCCCGGGAGGAGGAT  
CGTGTTCAGAGAAGCTCAACGGCAGCGTCTGCTGGCTCGGGCGCCCTGCCAGGCTGTTGTGCCGCTTGCA  
GGCCTTTGGGGAGACGCTGCTACTAGAGCTGGAGCAGGACTCCGGTGTGCAGGTGAGGGGGCTGACAGTGCAGTA  
CCTGGGCCAGGCGCCTGAGCTGCTGGGTGGAGCAGAGCCTGGCACCTACCTGACTGGCACCATCAATGGAGATCC  
GGAGTCGGTGGCATCTCTGCACTGGGATGGGGGAGCCCTGTTAGGCGTGTACAAATATCGGGGGGGCTGAACTCCA  
CCTCCAGCCCCGGAGGGAGGCCACCCCTAACTCTGCTGGGGGACCTGGGGCTCACATCCTACGCCGGAAGAGTCC  
TGCCAGCGGTCAAGGTCCCATGTGCAACGTCAAGGCTCCTCTTGGAAAGCCCCAGCCCCAGACCCGAAGAGCCAA  
GCGCTTTGCTTCACTGAGTAGATTTGTGGAGACACTGGTGGTGGCAGATGACAAGATGGCCGCATTCCACGGTGC  
GGGGCTAAAGCGCTACCTGCTAACAGTGATGGCAGCAGCAGCCAAGGCCCTTCAAGCACCCAAGCATCCGCAATCC  
TGTCAGCTTGGTGGTGAAGTCCGCTAGTGATCCTGGGGTCAGGCGAGGAGGGGGCCCCAAGTGGGGGCCAGTGCTGC  
CCAGACCCCTGCGCAGCTTCTGTGCCCTGGCAGCGGGGCTCAACACCCCTGAGGACTCGGGCCCTGACCACCTTTGA  
CACAGCCATTCTGTTTACCGTCAGGACCTGTGTGGAGTCTCCACTTGCACACGCTGGGTATGGCTGATGTGGG  
CACCGTCTGTGACCCGGCTCGGAGCTGTGCCATTGTGGAGGATGATGGGCTCCAGTCAGCCTTCACTGCTGCTCA  
TGAAGTGGGTGATGTCTTCAACATGCTCCATGACAACCTCCAAGCCATGCATCAGTTTGAATGGGCCTTTGAGCAC  
CTCTCGCCATGTGATGGCCCTGTGATGGCTCATGTGGATCCTGAGGAGCCCTGGTCCCCCTGCAGTGCCCGCTT  
CATCACTGACTTCCCTGGACATGCTATGGGCACTGTCTCTTAGACAAACCAGAGGCTCCATTGCATCTGCCTGT  
GACTTTCCCTGGCAAGGACTATGATGCTGACCGCCAGTGCCAGCTGACCTTCCGGGCCGACTCACGCCATTGTCC  
ACAGCTGCCCGGCCCTGTGCTGCCCTCTGGTGTCTGGCCACCTCAATGGCCATGCCATGTGCCAGACCAACA  
CTCGCCCTGGGCCGATGGCACACCCCTGCGGGGCCGACAGGCTGCATGGGTGGTGGTGCCTCCAGTGGACCA  
GCTCCAGGACTTCAATATTCACAGGCTGGTGGCTGGGGTCTTGGGGACCATGGGGTGAAGTGTCTCGGACCTG  
TGGGGGTGGTGTCCAGTTCTCCTCCGAGACTGCACGAGGCTGTCCCCCGGAATGGTGGCAAGTACTGTGAGGG  
CCGCCGTACCCGCTTCCGCTCCTGCAACACTGAGGACTGCCCAACTGGCTCAGCCCTGACCTTCCGCGAGGAGCA  
GTGTGCTGCCATACAACCACCGCACCGACCTCTTCAAGAGCTTCCAGGGGCCATGGACTGGGTTCCTCGCTACAC  
AGGCGTGGCCCCCAGGACCAAGTGCACACTGCCAGGCCGCGGCACTGGGCTACTACTATGTGCTGGAGCC  
ACGGGTGGTAGATGGGACCCCCCTGTTCCCCCGACAGCTCCTCGGTCTGTGTCCAGGGCCGATGCATCCATGGTGG  
CTGTGATCGCATCATTTGGCTCCAAGAAGAAGTTTGACAAGTGCATGGTGTGCGGAGGGGACGGTCTGGTTGCAG  
CAAGCAGTCAGGCTCCTTCAGGAAATTGAGGTACGGATACAACAATGTGGTCACTATCCCCGCGGGGGCCACCCA  
CATTCCTTGTCCGGCAGCAGGGAAACCCTGGCCACCGGAGCATCTACTTGGCCCTGAAGCTGCCAGATGGCTCCTA  
TGCCCTCAATGGTGAATACACGCTGATGCCCTCCCCACAGATGTGGTACTGCCTGGGGCAGTCAGCTTGGCGTA  
CAGCGGGGCACTGCAGCCTCAGAGACACTGTACGCCATGGGCCACTGGCCAGCCTTTGACACTGCAAGTCTCT  
AGTGGCTGGCAACCCCCAGGACACACGCTCCGATACAGCTTCTTCTGCCCCGGCCGACCCCTTCAACGCCACG  
CCCCACTCCCCAGGACTGGCTGCACCGAAGAGCACAGATTCTGGAGATCCTTCGGCGGCGCCCCCTGGGCGGGCAG  
GAA**ATA**ACCTCACTATCCCGGCTGCCCTTTCTGGGCACCGGGGCTCGGACTTAGCTGGGAGAAAGAGAGAGCTT  
CTGTTGCTGCCATGCTAAGACTCAGTGGGGAGGGGCTGTGGGCGTGAGACCTGCCCTCCTCTCTGCCCTAAT  
CGCGAGGCTGGCCCTGCCCTGGTTTCTGCCCTGGGAGGCAGTGATGGGTAGTGGATGGAAGGGGCTGACAGAC  
AGCCCTCCATCTAAACTGCCCCCTCTGCCCTGCGGTCACAGGAGGGAGGGGGAAGGCAGGGAGGGGCTGGGGC  
CAGTTGTATTTATTTAGTATTTATTCACTTTTATTTAGCACAGGGAAGGGGACAAGGACTAGGGTCTGGGGAA  
CCTGACCCCTGACCCCTCATAGCCCTCACCTGGGGCTAGGAAATCCAGGGTGGTGGTGTAGGTATAAGTGGTG  
TGTGTATGCGTGTGTGTGTGTGTGAAAATGTGTGTGTGCTTATGTATGAGGTACAACCTGTTCTGCTTTCTCTC  
TTCTGAATTTTATTTTGGGAAAAGAAAAGTCAAGGGTAGGGTGGGCCTTCAAGGAGTGAGGGATTATCTTTT  
TTTTTTTTTCTTTCTTTCTTTCTTTTTTTTTTTTGGAGACAGAATCTCGCTCTGTGCCCCAGGCTGGAGTGCAATG  
GCACAATCTCGGCTCACTGCATCCTCCGCTCCCGGTTCAAGTGATTCTCATGCCCTCAGCCTCCTGAGTAGCTG  
GGATTACAGGCTCCTGCCACCACGCCAGCTAATTTTTGTTTTGTTTTGTTTGGAGACAGAGTCTCGCTATTGTC  
ACCAGGGCTGGAATGATTTCACTCACTGCAACCTTCGCCACCTGGGTTCAGCAATTCTCCTGCCTCAGCCTCC  
CGAGTAGCTGAGATTATAGGCACCTACCACCACGCCCGGCTAATTTTTGTATTTTTTAGTAGAGACGGGGTTTCAC  
CATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTTAGGTATCCACTCGCCTTCATCTCCCAAAGTGCTGGGATT  
ACAGGCGTGAGCCACCGTGCCTGGCCACGCCAACTAATTTTTGTATTTTTTAGTAGAGACAGGGTTTACCATTGT  
TGGCCAGGCTGCTCTTGAACCTCTGACCTCAGGTATCGACCTGCCCTCGGCTCCCAAAGTGCTGGGATTACAGG  
TGTGAGCCACCACGCCCGGTACATATTTTTTAAATTGAATTTCTACTATTTATGTGATCCTTTTGGAGTCAGACAG



**FIGURE 180**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73492

><subunit 1 of 1, 837 aa, 1 stop

><MW: 90167, pI: 8.39, NX(S/T): 1

MSQTGSHPGRGLAGRWLWGAQPCLLLPIVPLSWLVWLLLLLLLASLLPSARLASPLPREEEIV  
FPEKLNGSVLPGSGAPARLLCRLQAFGETLLLELEQDSGVQVEGLTVQYLGQAPELLGGAEP  
GTYLTGTINGDPESVASLHWDGGALLGVLYRGAELHLQPLEGGTPNSAGGPGAHLRRKSP  
ASGQGPCMNVKAPLGSPSPRRRAKRFASLSRFVETLVVADDKMAAFHGAGLKRYLLTVMAA  
AAKAFKHPSIRNPVSLVVTRLVILGSGEEGPQVGPSAAQTLRSFCAWQRLNTPEDSGPDHF  
DTAILFTRQDLGCVSTCDTLGMADVGTVCDPARSCAIVEDDGLQSAFTAHELGHVFNMLHD  
NSKPCISLNGPLSTSRHVMAVMAHVDPEEPWSPCSARFITDFLDNGYGHCLLDKPEAPLHL  
PVTFFPGKDYDADRQCQLTFGPDSRHCPQLPPPCAALWCSGHLNGHAMCQTKHSPWADGTPCG  
PAQACMGGRCLHMDQLQDFNIPQAGGWGPWGPWGDCSRTC GGGVQFSSRDCTRVPVRNGGKY  
CEGRRTFRSCNTEDCPTGSALTFREEQCAAYNHRTDLFKSFPGPMDWVPRYTGVPAPQDQCK  
LTCQARALGYYYVLEPRVVDGTPCSPDSSSVQVQGRCIHAGCDRIIGSKKKFDKCMVCGGDG  
SGCSKQSGSFRKFRYGYNVVTIPAGATHILVRQQGNPGHRSIYLALKLPDGSYALNGEYTL  
MPSPTDVVLPGAVSLRYSGATAASETLSGHGPLAQPLTLQVLVAGNPQDTRLRYSFFVPRPT  
PSTPRPTQDWLHRRRAQILEILRRRPWAGRK

**Important features of the protein:**

**Signal peptide:**

amino acids 1-48

**N-glycosylation site.**

amino acids 68-71

**Glycosaminoglycan attachment site**

amino acids 188-191, 772-775

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 182-185

**Tyrosine kinase phosphorylation site.**

amino acids 730-736

**N-myristoylation sites.**

amino acids 5-10, 19-24, 121-126, 125-130, 130-135, 147-152, 167-  
172, 168-173, 174-179, 323-328, 352-357, 539-544, 555-560, 577-  
582, 679-684, 682-687, 763-768

**Amidation sites.**

amino acids 560-563, 834-837

**Leucine zipper pattern.**

amino acids 17-38, 24-45

**Neutral zinc metallopeptidases, zinc-binding region signature.**

amino acids 358-367

## FIGURE 181

CAGCAGTGGTCTCTCAGTCCTCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACC**ATGG**  
CAAAGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCC  
AAGAAAATATGTAAATCACTTAAGATTTGTGGACTGGTGTGTTGGTATCCTGGCCCTAACTCT  
AATTGTCCTGTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAAGCCTATGACA  
TGGAGCACACTTTCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTG  
ACCAGAACTGAAATATTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTT  
TAAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGA  
TTAAAGTGATTCTCTGAATTTTCTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACC  
ACAACTTTCTTTGAACAGTCAGTGATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGA  
TTTTCTTAAAAATTCCAAAATTCTGGAGATTTGTGATAACGTGACCATGTATTGGATCAATC  
CCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAGGAGGAGGGAGAAGATCTTCACTTT  
CCTGCCAACGAAAAAAAAAGGGATTGAACAAAATGAACAGTGGGTGGTCCCTCAAGTGAAAGT  
AGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGAAGAACTTCCAATAAATGACTATACTG  
AAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTATTTACTGCCGT  
CGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTTACTAGGCTACTACCCATATCCATA  
CTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTGTAACCTGGTGGGTGGCC  
GCATGCTGGGGAGGGTCT**TAA**TAGGAGGTTTGAGCTCAAATGCTTAAACTGCTGGCAACATAT  
AATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCCTGGTAGCCAGCT  
CTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACATTATCACC  
AAAAAAAAAAAAAAAAAAAAA

**FIGURE 182**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73727
><subunit 1 of 1, 317 aa, 1 stop
><MW: 37130, pI: 5.18, NX(S/T): 3
MAKNPPENCEDCHILNAEAFKSKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPPKAY
DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKT
QIKVIPEFSEPEEEIDENEEITTTFFEQSVIWWPAEKPIENRDFLKNKILEICDNVTMYWI
NPTLISVSELQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDY
TENGIEFDPMLEDERGYCCIIYCRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWV
ARMLGRV
```

**Important features of the protein:****Signal peptide:**

amino acids 1-40

**Transmembrane domain:**

amino acids 25-47 (type II)

**N-glycosylation sites.**

amino acids 94-97, 180-183

**Glycosaminoglycan attachment sites.**

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

**N-myristoylation sites.**

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

**Microbodies C-terminal targeting signal.**

amino acids 315-317

**Cytochrome c family heme-binding site signature.**

amino acids 9-14

**FIGURE 183**

GCGGAAGTGGCTCCGGCTGGCACCTGAGGAGCGGCGTGACCCCGAGGGCCCAGGGAGCTGCC  
CGGCTGGCCTAGGCAGGCAGCCGCACC**ATG**GCCAGCACGGCCGTGCAGCTTCTGGGCTTCCT  
GCTCAGCTTCCTGGGCATGGTGGGCACGTTGATCACCACCATCCTGCCGCACTGGCGGAGGA  
CAGCGCACGTGGGCACCAACATCCTCACGGCCGTGTCTACCTGAAAGGGCTCTGGATGGAG  
TGTGTGTGGCACAGCACAGGCATCTACCAGTGCCAGATCTACCGATCCCTGCTGGCGCTGCC  
CCAAGACCTCCAGGCTGCCCCGCGCCCTCATGGTCATCTCCTGCCTGCTCTCGGGCATAGCCT  
GCGCCTGCGCCGTTCATCGGGATGAAGTGACGCGCTGCGCCAAGGGCACACCCGCCAAGACC  
ACCTTTGCCATCCTCGGCGGCACCCTCTTCATCCTGGCCGGCCTCCTGTGCATGGTGGCCGT  
CTCCTGGACCACCAACGACGTGGTGCAGAACTTCTACAACCCGCTGCTGCCCAGCGGCATGA  
AGTTTGAGATTGGCCAGGCCCTGTACCTGGGCTTCATCTCCTCGTCCCTCTCGCTCATTGGT  
GGCACCCTGCTTTGCCTGTCCTGCCAGGACGAGGCACCCTACAGGCCCTACCAGGCCCCGCC  
CAGGGCCACCACGACCACTGCAAACACCGCACCTGCCTACCAGCCACCAGCTGCCTACAAAG  
ACAATCGGGCCCCCTCAGTGACCTCGGCCACGCACAGCGGGTACAGGCTGAACGACTACGTG  
**TGAGT**CCCCACAGCCTGCTTCTCCCCTGGGCTGCTGTGGGCTGGGTCCCCGGCGGGACTGTC  
AATGGAGGCAGGGGTTCAGCACAAAGTTTACTTCTGGGCAATTTTTGTATCCAAGGAAATA  
ATGTGAATGCGAGGAAATGTCTTTAGAGCACAGGGACAGAGGGGGAAATAAGAGGAGGAGAA  
AGCTCTCTATACCAAAGACTGAAAAAAAAAATCCTGTCTGTTTTTGTATTTATTATATATAT  
TTATGTGGGTGATTTGATAACAAGTTTAATATAAAGTGACTTGGGAGTTTGGTCAGTGGGGT  
TGGTTTGTGATCCAGGAATAAACCTTGCGGATGTGGCTGTTTATGAAAAAAAAAAAAA

## **FIGURE 184**

MASTAVQLLGFLLSFLGMVGTLITTLPHWRRTAHVGTNILTAVSYLKGLWMECVWHSTGIY  
QCQIYRSLALPQDLQAARALMVISCLLSGIACACAVIGMKCTRCAKGTPAKTTFAILGGTL  
FILAGLLCMVAVSWTTNDVVQNFYNPLLPSPGMKFEIGQALYLGFISSSLSLIGGTLLCLSCQ  
DEAPYRPHYQAPPRATTTTANTAPAYQPPAAYKDNRAPSVTSATHSGYRLNDYV

**Important features of the protein:**

**Signal peptide:**

amino acids 1-21

**Transmembrane domains:**

amino acids 82-103, 115-141, 160-182

**FIGURE 185**

GAGCTCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCG  
 CAGGCGGCAGGGCGGGCGGCCAGGATC**ATG**TCCACCACCACATGCCAAGTGGTGGCGTTCTCT  
 CCTGTCCATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCC  
 AGGACCTGTACGACAACCCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGC  
 GTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCCCTATTTACCATCCTGGGACTTCCAGC  
 CATGCTGCAGGCAGTGCGAGCCCTGATGATCGTAGGCATCGTCCTGGGTGCCATTGGCCTCC  
 TGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCC  
 AACATGACACTGACCTCCGGGATCATGTTCAATTGTCTCAGGTCTTTGTGCAATTGCTGGAGT  
 GTCTGTGTTTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCG  
 GCATGGGTGGGATGGTGCAGACTGTTCCAGACCAGGTACACATTTGGTGCGGCTCTGTTTCGTG  
 GGCTGGGTTCGCTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGG  
 CCTGGCACCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTTG  
 CCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAAACAAGAAG  
 ATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCAAGCACGACTA  
 TGTG**TAA**TGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCACCCAAAA  
 AACAAGGAGATCCCATCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGC  
 CTCGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTTCC  
 ACCATAAAACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTAT  
 TTCTTTTTTTAAATATAACTTTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCAC  
 ATTTTGATGATTTAGACAGACTCCCCCTCTTCTCCTAGTCAATAAACCCATTGATGATCTA  
 TTTCCCAGCTTATCCCCAAGAAAACCTTTTGAAAGGAAAGAGTAGACCCAAAGATGTTATTTT  
 CTGCTGTTTGAATTTTGTCTCCCCACCCCAACTTGGCTAGTAATAAACACTTACTGAAGAA  
 GAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTG  
 TGATCTTAAAGTTACCAAACCAAAGTCATTTTCAGTTTGAGGCAACCAAACCTTTCTACTG  
 CTGTTGACATCTTCTTATTACAGCAACACCATTTCTAGGAGTTTCCTGAGCTCTCCACTGGAG  
 TCCTCTTTCTGTGCGGGTCAGAAATTGTCCTTAGATGAATGAGAAAATTATTTTTTTTAAAT  
 TTAAGTCCTAAATATAGTTAAAATAAATAATGTTTTAGTAAATGATACACTATCTCTGTGA  
 AATAGCCTCACCCCTACATGTGGATAGAAGGAAATGAAAAAATAATTGCTTTGACATTGTCT  
 ATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTCCATGAAAAGCTCACACCTGTAATC  
 CTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGTTCGAGACTAGCCTG  
 GGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCATGGTG  
 GCATACACCTGTAGTCCCAGCATTCGGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGG  
 AGGTTGGGGCTGCAGTGAGCCATGATCACACCACTGCACTCCAGCCAGGTGACATAGCGAGA  
 TCCTGTCTAAAAAAATAAAAAATAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAA  
 ACTAATTCTTTAA



**FIGURE 186**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73734

><subunit 1 of 1, 261 aa, 1 stop

><MW: 27856, pI: 8.50, NX(S/T): 1

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVVFQYEGWLWRSCVRQSSGFTE  
CRPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIM  
FIVSGLCAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTL  
IGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTE  
DEVQSYPSKHDYV

**Signal peptide:**

amino acids 1-23

**Transmembrane domains:**

amino acids 81-100, 121-141, 173-194

**FIGURE 187**

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCG  
GAGTCCAGCTGGCTAAAACTCATCCCAGAGGATA**ATG**GCAACCCATGCCTTAGAAATCGCTG  
GGCTGTTTCTTGGTGGTGGTGGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGG  
AGAGTGTCGGCCTTCATTGAAAACAACATCGTGGTTTTTGAACCTTCTGGGAAGGACTGTG  
GATGAATTGCGTGAGGCAGGCTAACATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGG  
CTCTTTCTCCGGACCTACAGGCAGCCAGAGGACTGATGTGTGCTGCTTCCGTGATGTCCTTC  
TTGGCTTTCATGATGGCCATCCTTGGCATGAAATGCACCAGGTGCACGGGGGACAATGAGAA  
GGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCATCATCACGGGCATGGTGGTGC  
TCATCCCTGTGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTATAACTCAATAGTGAAT  
GTTGCCCAAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCACTGGTGCT  
GATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTACA  
GATACTCGATACCTTCCCATCGCACAAACCCAAAAAAGTTATCACACCGGAAAGAAGTCACCG  
AGCGTCTACTCCAGAAGTCAGTATGTG**TAG**TTGTGTATGTTTTTTAACTTTACTATAAAGC  
CATGCAAATGACAAAAATCTATATTACTTTCTCAAATGGACCCCAAAGAACTTTGATTTA  
CTGTTCTTAACTGCCTAATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAA  
GCTATTTTCAGCAGAATGAGATATTAAACCCAATGCTTTGATTGTTCTAGAAAGTATAGTAAT  
TTGTTTTCTAAGGTGGTTCAAGCATCTACTCTTTTTTATCATTTACTTCAAATGACATTGCT  
AAAGACTGCATTATTTTACTACTGTAATTTCTCCACGACATAGCATTATGTACATAGATGAG  
TGTAACATTTATATCTCACATAGAGACATGCTTATATGGTTTTTATTTAAATGAAATGCCAG  
TCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAAATCATGGATAGGGTTG  
AAGAAGGTTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCCTCCATTTATAATGA  
AGATTAAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATATGCTG  
TTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCCAGAGGCTTTTTTT  
TTCTTGTGTATTAAATTAACATTTTTTAAACGCAGATATTTTGTCAAGGGGCTTTCATTCA  
AACTGCTTTTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATG  
GTTTTAGGAAAGTGAAAATATTTTTGTTTTTGTATTTGAAGAAGAATGATGCATTTTGACAA  
GAAATCATATATGTATGGATATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATC  
AATATAAATAAAAGAGCAGAAAAATATGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACA  
ACAAAAAAGTTGTCCTTTGAGAACTTCACCTGCTCCTATGTGGGTACCTGAGTCAAAATTG  
TCATTTTTGTTCTGTGAAAAATAAATTTCTTCTGTACCATTCTGTTTAGTTTTACTAAA  
ATCTGTAAATACTGTATTTTTCTGTTTATTCCAAATTTGATGAACTGACAATCCAATTTGA  
AAGTTTGTGTGACGTCTGTCTAGCTTAAATGAATGTGTTCTATTTGCTTTATACATTTATA  
TTAATAAATTGTACATTTTTCTAATT

**FIGURE 188**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73735

><subunit 1 of 1, 225 aa, 1 stop

><MW: 24845, pI: 9.07, NX(S/T): 0

MATHALEIAGLFLGGVGMVGTVAVTVMPPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRM  
QCKIYDSLALLSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGI  
IFIITGMVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIIVGGALFCCVF  
CCNEKSSSYRYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

**Signal peptide:**

amino acids 1-17

**Transmembrane domains:**

amino acids 82-101, 118-145, 164-188

**FIGURE 189**

TCGCC**ATG**GCCTCTGCCGGAATGCAGATCCTGGGAGTCGTCCTGACACTGCTGGGCTGGGTG  
AATGGCCTGGTCTCCTGTGCCCTGCCCATGTGGAAGGTGACCGCTTTCATCGGCAACAGCAT  
CGTGGTGGCCCAGGTGGTGTGGGAGGGCCTGTGGATGTCCTGCGTGGTGCAGAGCACCGGCC  
AGATGCAGTGCAAGGTGTACGACTCACTGCTGGCGCTGCCACAGGACCTGCAGGCTGCACGT  
GCCCTCTGTGTCATCGCCCTCCTTGTGGCCCTGTTTCGGCTTGCTGGTCTACCTTGCTGGGGC  
CAAGTGTACCACCTGTGTGGAGGAGAAGGATTCCAAGGCCCGCCTGGTGCTCACCTCTGGGA  
TTGTCTTTGTCATCTCAGGGGTCCTGACGCTAATCCCCGTGTGCTGGACGGCGCATGCCATC  
ATCCGGGACTTCTATAACCCCCCTGGTGGCTGAGGCCCAAAGCGGGAGCTGGGGGCCTCCCT  
CTACTTGGGCTGGGCGGCCTCAGGCCTTTTGTGCTGGGTGGGGGGTTGCTGTGCTGCACTT  
GCCCCCTCGGGGGGGTCCCAGGGCCCCAGCCATTACATGGCCCGCTACTCAACATCTGCCCCCT  
GCCATCTCTCGGGGGCCCTCTGAGTACCCTACCAAGAATTACGTC**TGA**CGTGGAGGGGAATG  
GGGGCTCCGCTGGCGCTAGAGCCATCCAGAAGTGGCAGTGCCCAACAGCTTTGGGATGGGTT  
CGTACCTTTTGTCTCTGCCTCCTGCTATTTTTCTTTTGAAGGATATTTAAATTCATTT  
GAAACTGAGCCAAGGTGTTGACTCAGACTCTCACTTAGGCTCTGCTGTTTCTCACCTTGG  
ATGATGGAGCCAAAGAGGGGATGCTTTGAGATTCTGGATCTTGACATGCCCATCTTAGAAGC  
CAGTCAAGCTATGGAATAATGCGGAGGCTGCTTGCTGTGCTGGCTTTGCAACAAGACAGAC  
TGTCCCCAAGAGTTCCTGCTGCTGCTGGGGGCTGGGCTTCCCTAGATGTCACTGGACAGCTG  
CCCCCATCCTACTCAGGTCTCTGGAGCTCCTCTCTTACCCCTGGAAAAACAAATCATCTG  
TTAACAAGGACTGCCACCTCCGGAATTTCTGACCTCTGTTTCCTCCGTCCTGATAAGACG  
TCCACCCCCCAGGGCCAGGTCCCAGCTATGTAGACCCCCGCCCCACCTCCAACACTGCACC  
CTTCTGCCCTGCCCCCTCGTCTCACCCCTTTACACTCACATTTTTTATCAAATAAAGCATG  
TTTTGTAGTGCA

**FIGURE 190**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73736
><subunit 1 of 1, 220 aa, 1 stop
><MW: 23292, pI: 8.43, NX(S/T): 0
MASAGMQILGVVLTLLGWVNGLVSCALPMWKVTAFIGNSIVVAQVVWEGLWMSCVVQSTGQM
QCKVYDSLLALPQDLQAARALCVIALLVALLFGLLVYLAGAKCTTCVEEKDSKARLVLTSGIV
FVISGVLTLIPVCWTAHAIIRDFYNPLVAEAQKRELGASLYLGWAASGLLLLGGGLLCCTCP
SGGSQGPPSHYMARYSTSAPAIRGPSEYPTKNYV
```

**Transmembrane domains:**

amino acids 8-30 (type II), 82-102, 121-140, 166-186

**FIGURE 191**

GCCAAGGAGAACATCATCAAAGACTTCTCTAGACTCAAAGGCTTCCACGTTCTACATCTTG  
AGCATCTTCTACCACTCCGAATTGAACCAGTCTTCAAAGTAAAGGCAATGGCATTATCC  
TTGCAAATTGCTGGGCTGGTTCCTGGGTTCCCTGGCATGGTGGGACTCTTGCCACAACCCT  
TCTGCCTCAGTGGTGGAGTATCAGCTTTTGTGGCAGCAACATTATTGTCTTTGAGAGGCTC  
TGGGAAGGGCTCTGGATGAAATTGCATCCGACAAGCCAGGGTCCGGTTGCAATGCAAGTTCTA  
TAGTCCTTGTTGGCTCTCCCGCCTGCCCTGGAAACAGCCCGGGCCCTCATGTGTGTGGCTG  
TTGCTCTCTCCTTGATCGCCCTGCTTATTGGCATCTGTGGCATGAAGCAGGTCCAGTGCACA  
GGCTCTAACGAGAGGGCCAAAGCATACTTCTGGGAACCTTCAGGAGTCCTCTTCATCCTGAC  
GGGTATCTTCGTTCTGATTCCGGTGAGCTGGACAGCCAATATAATCATCAGAGATTTCTACA  
ACCCAGCCATCCACATAGGTCAGAAACGAGAGCTGGGAGCAGCACTTTTCCTTGGCTGGGCA  
AGCGCTGCTGTCCTCTTCATTGGAGGGGGTCTGCTTTGTGGATTTTGCTGCTGCAACAGAAA  
GAAGCAAGGGTACAGATATCCAGTGCCTGGCTACCGTGTGCCACACACAGATAAGCGAAGAA  
ATACGACAATGCTTAGTAAGACCTCCACCAGTTATGTCTTAATGCCTCCTTTTGGCTCCAAGT  
ATGGACTATGGTCAATGTTTTTTATAAAGTCCTGCTAGAACTGTAAGTATGTGAGGCAGGA  
GAACTTGCTTTATGTCTAGATTTACATTGATACGAAAGTTTCAATTTGTTACTGGTGGTAGG  
AATGAAAATGACTTACTTGGACATTCTGACTTCAGGTGTATTAAATGCATTGACTATTGTTG  
GACCCAATCGCTGCTCCAATTTTCATATTCTAAATTCAAGTATACCCATAATCATTAGCAAG  
TGTACAATGATGGACTACTTATTACTTTTTGACCATCATGTATTATCTGATAAGAATCTAAA  
GTTGAAATTGATATTCTATAACAATAAAACATATACCTATTCTA

**FIGURE 192**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73737
><subunit 1 of 1, 173 aa, 1 stop
><MW: 18938, pI: 9.99, NX(S/T): 1
MNCIRQARVRLQCKFYSSLLALPPALETARALMCVAVALSLIALLLIGICGMKQVQCTGSNER
AKAYLLGTSGVLFILTGIFVLIPVSWTANIIIRDFYNPAIHIGQKRELGAALFLGWASAAVL
FIGGGLLCGFCCCNRRKKQGYRYPVPGYRVPHTDKRRNTTMLSKTSTSYV
```

**Important features of the protein:****Transmembrane domains:**

amino acids 31-51, 71-90, 112-133

**N-glycosylation site.**

amino acids 161-164

### FIGURE 193

AGTGACAATCTCAGAGCAGCTTCTACACCACAGCCATTTCCAGCATGAAGATCACTGGGGGT  
CTCCTTCTGCTCTGTACAGTGGTCTATTTCTGTAGCAGCTCAGAAAGCTGCTAGTCTGTCTCC  
AAAAAAAGTGGACTGCAGCATTTACAAGAAGTATCCAGTGGTGGCCATCCCCTGCCCATCA  
CATACCTACCAGTTTGTGGTCTGACTACATCACCTATGGGAATGAATGTCACTTGTGTACC  
GAGAGCTTGAAAAGTAATGGAAGAGTTCAGTTTCTTCACGATGGAAGTTGCTAAATTCTCCA  
TGGACATAGAGAGAAAGGAATGATATTCTCATCATCATCTTCATCATCCCAGGCTCTGACTG  
AGTTTCTTTCAGTTTTACTGATGTTCTGGGTGGGGGACAGAGCCAGATTCAGAGTAATCTTG  
ACTGAATGGAGAAAGTTTCTGTGCTACCCCTACAAACCCATGCCTCACTGACAGACCAGCAT  
TTTTTTTTTAAACACGTCAATAAAAAAATAATCTCCCAGA



## FIGURE 194

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73739
```

><subunit 1 of 1, 85 aa, 1 stop

><MW: 9232, pI: 7.94, NX(S/T): 0

MKITGGLLLLCTVVYFCSSSEAAASLSPKKVDCSIYKKYPVVAIPCPITYLPVCGSDYITYGN  
 ECHLCTESLKSNGRVQFLHDGSC

**Signal peptide:**

amino acids 1-19

**FIGURE 195**

CCCGCGCCCGGTTCTCCCTCGCAGCACCTCGAAGTGCGCCCTCGCCCTCCTGCTCGCGCCC  
CGCCGCC**ATG**GCTGCCTCCCCGCGCGGCCTGCTGTCCTGGCCCTGACCGGGCTGGCGCTGC  
TCCTGCTCCTGTGCTGGGGCCCAGGTGGCATAAGTGGAAATAAACTCAAGCTGATGCTTCAA  
AAACGAGAAGCACCTGTTCCAACTAAGACTAAAGTGGCCGTTGATGAGAATAAAGCCAAAGA  
ATTCTTGGCAGCCTGAAGCGCCAGAAGCGGCAGCTGTGGGACCGGACTCGGCCCCGAGGTGC  
AGCAGTGGTACCAGCAGTTTCTCTACATGGGCTTTGATGAAGCGAAATTTGAAGATGACATC  
ACCTATTGGCTTAACAGAGATCGAAATGGACATGAATACTATGGCGATTACTACCAACGTCA  
CTATGATGAAGACTCTGCAATTGGTCCCCGGAGCCCCCTACGGCTTTAGGCATGGAGCCAGCG  
TCAACTACGATGACTAC**TAA**CCATGACTTGCCACACGCTGTACAAGAAGCAAATAGCGATTC  
TCTTCATGTATCTCCTAATGCCTTACACTACTTGGTTTCTGATTTGCTCTATTTTCAGCAGAT  
CTTTTCTACCTACTTTGTGTGATCAAAAAGAAGAGTTAAAACAACACATGTAAATGCCTTT  
TGATATTTTCATGGGAATGCCTCTCATTTAAAAATAGAAATAAAGCATTTTGTTAAAAAGA

### FIGURE 196

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73742
```

```
><subunit 1 of 1, 148 aa, 1 stop
```

><MW: 17183, pI: 8.77, NX(S/T): 0

MAASPARPAVLALTGLALLLLLCWGPGGISGNKLKMLQKREAPVPTKTKVAVDENKAKEFL  
GSLKRQKRQLWDRTRPEVQQWYQQFLYMGFDEAKFEDDITYWLNDRNGHEYYGDYYQRHYD  
EDSAIGPRSPYGFRHGASVNYDDY

Signal peptide:

amino acids 1-30

**FIGURE 197**

CGGCTCGAGCCCGCCCGGAAGTGCCCGAGGGGCGCGATGGAGCTGGGGGAGCCGGGCGCTC  
GGTAGCGCGGCGGGCAAGGCAGGCGCC**ATG**ACCCTGATTGAAGGGGTGGGTGATGAGGTGAC  
CGTCCTTTTCTCGGTGCTTGCCTGCCTTCTGGTGCTGGCCCTTGCCTGGGTCTCAACGCACA  
CCGCTGAGGGCGGGGACCCACTGCCCCAGCCGTCAGGGACCCCAACGCCATCCCAGCCCAGC  
GCAGCCATGGCAGCTACCGACAGCATGAGAGGGGAGGCCCCAGGGGCAGAGACCCCCAGCCT  
GAGACACAGAGGTCAAGCTGCACAGCCAGAGCCCAGCACGGGGTTCACAGCAACACCGCCAG  
CCCCGGA CTCCCCGCAGGAGCCCCCTCGTGCTACGGCTGAAATTCTCAATGATTCAGAGCAG  
GTGGCCAGGGCCTGGCCCCACGACACCATTGGCTCCTTGAAAAGGACCCAGTTTCCCGGCCG  
GGAACAGCAGGTGCGACTCATCTACCAAGGGCAGCTGCTAGGCGACGACACCCAGACCCTGG  
GCAGCCTTCACCTCCCTCCCAACTGCGTTCTCCACTGCCACGTGTCCACGAGAGTCGGTCCC  
CCAAATCCCCCCTGCCCGCCGGGGTCCGAGCCCGGCCCTCCGGGCTGGAAATCGGCAGCCT  
GCTGCTGCCCCTGCTGCTCCTGCTGTTGCTGCTGCTCTGGTACTGCCAGATCCAGTACCGGC  
CCTTCTTTCCCCTGACCGCCACTCTGGGCCTGGCCGGCTTCACCCTGCTCCTCAGTCTCCTG  
GCCTTTGCCATGTACCGCCCG**TAG**TGCCTCCGCGGGCGCTTGGCAGCGTCGCCGGCCCCCTCC  
GGACCTTGCTCCCCGCGCCGCGGCGGGAGCTGCTGCCTGCCAGGCCCGCCTCTCCGGCCTG  
CCTCTTCCCGCTGCCCTGGAGCCCAGCCCTGCGCCGCAGAGGACTCCCGGGACTGGCGGAGG  
CCCCGCCCTGCGACCGCCGGGGCTCGGGGCCACCTCCCGGGGCTGCTGAACCTCAGCCC GCA  
CTGGGAGTGGGCTCCTCGGGGTCGGGCATCTGCTGTCGCTGCCTCGGCCCCGGGCAGAGCCG  
GGCCGCCCCGGGGGCCCCGTCTTAGTGTTCTGCCGGAGGACCCAGCCGCTCCAATCCCTGAC  
AGCTCCTTGGGCTGAGTTGGGGACGCCAGGTCGGTGGGAGGCTGGTGAAGGGGAGCGGGGAG  
GGGCAGAGGAGTTCCCCGGAACCCGTGCAGATTAAAGTAACTGTGAAGTTTTAAAAAAAAA  
AAAAAAA

**FIGURE 198**

MTLIEGVGDEVTVLFSVLACLLVLALAWVSTHTAEGGDPLPQPSGTPTPSQPSAAMAATDSM  
RGEAPGAETPSLRHRGQAAQPEPSTGFTATPPAPDSPQEPLVLRLKFLNDSEQVARAWPHDT  
IGSLKRTQFPGREQQVRLIYQGQLLGDDTQTLGSLHLPPNCVLHCHVSTRVGPPNPPCPPGS  
EPGPSGLEIGSLLLPLLLLLLLLLLWYCQIQYRFFFLPTATLGLAGFTLLLSLLAFAMYRP

Signal peptide:

amino acids 1-31

**Transmembrane domain:**

amino acids 195-217



```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73746
><subunit 1 of 1, 148 aa, 1 stop
><MW: 16896, pI: 6.05, NX(S/T): 1
MTKALLIYLVSSFLALNQASLISRCDLAQVLQLEDLDGFEGYSLSDWLCLAFVESKFNISKI
NENADGSFDYGLFQINSHYWCNDYKSYSENLC HVDCQDLLNP NLLAGIHC AKRIVSGARGMN
NWVEWRLHCSGRPLSYWLTGCRLR
```

amino acids 1-18

**FIGURE 201**

TCTGACCTGACTGGAAGCGTCCAAAGAGGGACGGCTGTCAGCCCTGCTTGACTGAGAACCCA  
 CCAGCTCATCCCAGACACCTCATAGCAACCTATTTATACAAAGGGGGAAAGAAACACCTGAG  
 CAGAATGGAATCATTATTTTTTTTCCCAAGGAGAAAACCGGGGTAAAGGGAGGGAAGCAATTC  
 AATTTGAAGTCCCTGTGAATGGGCTTTTCAAGGCAATTAAAGAAATCCACTCAGAGAGGAC  
 TTGGGGTGAAACTTGGGTCTGTGGTTTTCTGATTGTAAGTGGAAGCAGGTCTTGCACACGC  
 TGTTGGCAAATGTCAGGACCAGGTAAAGTGAAGTGGCAGAAAACTTCCAGGTGGAACAAGCA  
 ACCCATGTTCTGCTGCAAGCTTGAAGGAGCCTGGAGCGGGAGAAAGCTAACTTGAACATGAC  
 CTGTTGCATTTGGCAAGTTCTAGCAAC**ATG**CTCCTAAGGAAGCGATACAGGCACAGACCATG  
 CAGACTCCAGTTCTCTGCTGCTCCTGATGCTGGGATGCGTCCTGATGATGGTGGCGATGT  
 TGCACCCTCCCCACCACACCCTGCACCAGACTGTCACAGCCCAAGCCAGCAAGCACAGCCCT  
 GAAGCCAGGTACCGCCTGGACTTTGGGGAATCCCAGGATTGGGTACTGGAAGCTGAGGATGA  
 GGGTGAAGAGTACAGCCCTCTGGAGGGCCTGCCACCCTTTATCTCACTGCGGGAGGATCAGC  
 TGCTGGTGGCCGTGGCCTTACCCCAGGCCAGAAGGAACCAGAGCCAGGGCAGGAGAGGTGGG  
 AGCTACCGCCTCATCAAGCAGCCAAGGAGGCAGGATAAGGAAGCCCCAAAGAGGGGACTGGGG  
 GGCTGATGAGGACGGGGAGGTGTCTGAAGAAGAGGAGTTGACCCCGTTCAGCCTGGACCCAC  
 GTGGCCTCCAGGAGGCACTCAGTGCCCGCATCCCCCTCCAGAGGGCTCTGCCCCAGGTGCGG  
 CACCCACTGTGTCTGCAGCAGCACCCCTCAGGACAGCCTGCCACAGCCAGCGTCATCCTCTG  
 TTTCCATGATGAGGCCTGGTCCACTCTCCTGCGGACTGTACACAGCATCCTCGACACAGTGC  
 CCAGGGCCTTCTGAAGGAGATCATCCTCGTGGACGACCTCAGCCAGCAAGGACAACCTCAAG  
 TCTGCTCTCAGCGAATATGTGGCCAGGCTGGAGGGGGTGAAGTTACTCAGGAGCAACAAGAG  
 GCTGGGTGCCATCAGGGGCCCGGATGCTGGGGGCCACCAGAGCCACCGGGGATGTGCTCGTCT  
 TCATGGATGCCCACTGCGAGTGCCACCCAGGCTGGCTGGAGCCCCCTCCTCAGCAGAATAGCT  
 GGTGACAGGAGCCGAGTGGTATCTCCGGTGATAGATGTGATTGACTGGAAGACTTTCCAGTA  
 TTACCCCTCAAAGGACCTGCAGCGTGGGGTGTGGACTGGAAGCTGGATTTCCACTGGGAAC  
 CTTTGCCAGAGCATGTGAGGAAGGCCCTCCAGTCCCCCATAAGCCCCATCAGGAGCCCTGTG  
 GTGCCCCGAGAGGTGGTGGCCATGGACAGACATTACTTCCAAAACACTGGAGCGTATGACTC  
 TCTTATGTCGCTGCGAGGTGGTGAAAACCTCGAACTGTCTTTCAAGGCCTGGCTCTGTGGTG  
 GCTCTGTTGAAATCCTTCCCTGCTCTCGGGTAGGACACATCTACCAAATCAGGATTCCCAT  
 TCCCCCTCGACCAGGAGGCCACCCTGAGGAACAGGGTTTCGCATTGCTGAGACCTGGCTGGG  
 GTCATTCAAAGAAACCTTCTACAAGCATAGCCCAGAGGCCTTCTCCTTGAGCAAGGCTGAGA  
 AGCCAGACTGCATGGAACGCTTGCAGCTGCAAAGGAGACTGGGTTGTTCGGACATTCCACTGG  
 TTTCTGGCTAATGTCTACCCTGAGCTGTACCCATCTGAACCCAGGCCAGTTTCTCTGGAAA  
 GCTCCACAACACTGGACTTGGGCTCTGTGCAGACTGCCAGGCAGAAGGGGACATCCTGGGCT  
 GTCCCATGGTGTGGCTCCTTGCAGTGACAGCCGGCAGCAACAGTACCTGCAGCACACCAGC  
 AGGAAGGAGATTCACTTTGGCAGCCCACAGCACCTGTGCTTTGCTGTCAGGCAGGAGCAGGT  
 GATTCTTCAGAACTGCACGGAGGAAGGCCTGGCCATCCACCAGCAGCACTGGGACTTCCAGG  
 AGAATGGGATGATTGTCCACATTCTTTCTGGGAAATGCATGGAAGCTGTGGTGCAAGAAAAC  
 AATAAAGATTTGTACCTGCGTCCGTGTGATGGAAAAGCCCGCCAGCAGTGGCGATTTGACCA  
 GATAAATGCTGTGGATGAACGA**TGA**ATGTCAATGTCAGAAGGAAAAGAGAATTTTGGCCATC  
 AAAATCCAGCTCCAAGTGAACGTAAAGAGCTTATATATTTTCATGAAGCTGATCCTTTTGTGT  
 GTGTGCTCCTTGTGTTAGGAGAGAAAAAGCTCTATGAAAGAATATAGGAAGTTTCTCCTTT  
 TCACACCTTATTTCAATTGACTGCTGGCTGCTTA



**FIGURE 202**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73760

><subunit 1 of 1, 639 aa, 1 stop

><MW: 73063, pI: 6.84, NX(S/T): 2

MLLRKRYRHRPCRLQFLLLLLMLGCVLMMVAMLHPPHHTLHQTVTAQASKHSPEARYRLDFG  
ESQDWVLEAEDEGEYSPLEGLPPFISLREDQLLVAVALPQARRNQSQGRRGGSYRLIKQPR  
RQDKEAPKRDWGADEDEGEVSEEEELTPFSLDPRGLQEALSARIPLQRALPEVRHPLCLQQHP  
QDSLPTASVILCFHDEAWSTLLRTVHSILDTVPRAFLKEIILVDDLSQQGQLKSALSEYVAR  
LEGVKLLRSNKRLGAIRARMLGATRATGDVLVFMDAHCECHPGWLEPLLSRIAGDRSRVVS  
VIDVIDWKTFQYYPSKDLQRGVLDWKLDHFHWEPLPEHVRKALQSPISPIRSPVVPGEVVAMD  
RHYFQNTGAYDSLMSLRGGENLELSFKAWLCGGSV EILPCSRVGHYQNDSDHSPLDQEATL  
RNRVRIAETWLGSFKETFYKHSPEAFSLSKAEKPDCMERLQLQRRLLGCRTFHWFLANVYPEL  
YPSEPRPSFSGKLHNTGLGLCADCQAEGDILGCPMVLAPCSDSRQQQYLQHTSRKEIHFGSP  
QHLCFAVRQEQVILQNCTEEGLAIHQQHWDFQENGMIVHILSGKCM EAVVQENNKDLYLRPC  
DGKARQQWRFDQINAVDER

**Signal peptide:**

amino acids 1-28

**FIGURE 203**

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

CGCCAAGCATGCAGTAAAGGCTGAAAATCTGGGTCACAGCTGAGGAAGACCTCAGAC**ATG**GGA  
 GTCCAGGATGTGGCCTGCGCTGCTGCTGTCCCACCTCCTCCCTCTCTGGCCACTGCTGTTGC  
 TGCCCCCTCCCACCGCCTGCTCAGGGCTCTTCATCCTCCCCCTCGAACCCCACCAGCCCCAGCC  
 CGCCCCCGGTGTGCCAGGGGAGGCCCCCTCGGCCCCACGTATGTGTGCGTGTGGGAGCGAGC  
 ACCTCCACCAAGCCGATCTCCTCGGGTCCCAAGATCACGTGCGCAAGTCCTGCCTGGCACTG  
 CACCCCCAGCCACCCCATCAGGCTTTGAGGAGGGGCGCCCTCATCCCAATACCCCTGGGCT  
 ATCGTGTGGGGTCCCACCGTGTCTCGAGAGGATGGAGGGGACCCCAACTCTGCCAATCCCGG  
 ATTTCTGGACTATGGTTTTTGCAGCCCCCTCATGGGCTCGCAACCCCACACCCCAACTCAGACT  
 CCATGCGAGGTGATGGAGATGGGCTTATCCTTGGAGAGGCACCTGCCACCCTGCGGCCATTC  
 CTGTTGCGGGGGCCGTGGGGAAGGTGTGGACCCCCAGCTCTATGTCACAATTACCATCTCCAT  
 CATCATTTGTTCTCGTGGCCACTGGCATCATCTTCAAGTTCTGCTGGGACCGCAGCCAGAAGC  
 GACGCAGACCCCTCAGGGCAGCAAGGTGCCCTGAGGCAGGAGGAGAGCCAGCAGCCACTGACA  
 GACCTGTCCCCGGCTGGAGTCACTGTGCTGGGGGCCCTTCGGGGACTCACCTACCCCCACCCC  
 TGACCATGAGGAGCCCCGAGGGGGGACCCCGGCCTGGGATGCCCCACCCCAAGGGGGCTCCAG  
 CCTTCCAGTTGAACCGGT**TGA**GGGCAGGGGCAATGGGATGGGAGGGCAAAGAGGGGAAGGCAAC  
 TTAGGTCTTCAGAGCTGGGGTGGGGGTGCCCTCTGGATGGGTAGTGAGGAGGCAGGCGTGGC  
 CTCCCACAGCCCCTGGCCCTCCCAAGGGGGCTGGACCAGCTCCTCTCTGGGAGGCACCCTTC  
 CTTCTCCCAGTCTCTCAGGATCTGTGTCTTAATTCTCTGCTGCCCATAACTCCAACCTCTGCCC  
 TCTTTGGTTTTTTTCTCATGCCACCTTGTCTAAGACAACCTCTGCCCTCTTAACCTTGATTCCC  
 CCTCTTTGTCTTGAACCTTCCCCCTTCTATTCTGGCCTACCCCTTGGTTCCTGACTGTGCCCTT  
 TCCCTCTTCCTCTCAGGATTCCCCTGGTGAATCTGTGATGCCCCCAATGTTGGGGTGCAGCC  
 AAGCAGGAGGCCAAGGGGGCCGGCACAGCCCCCATCCACTGAGGGTGGGGCAGCTGTGGGGA  
 GCTGGGGGCCACAGGGGCTCCTGGCTCCTGCCCCCTTGACACACCACCCGGAACACTCCCCAGCC  
 CCACGGGCAATCCTATCTGCTCGCCCTCCTGCAGGTGGGGGCCCTCACATATCTGTGACTTCG  
 GGTCCCTGTCCCCACCCTTGTGCACTCACATGAAAGCCTTGACACTCACCTCCACCTTCAC  
 AGGCCATTTGCACACGCTCCTGCACCCTCTCCCCGTCCATAACCGCTCCGCTCAGCTGACTCT  
 CATGTTCTCTCGTCTCACATTTGCACTCTCTCCTTCCCACATTCTGTGCTCAGCTCACTCAG  
 TGGTCAGCGTTTCCTGCACACTTTACCTCTCATGTGCGTTTCCCGGCCTGATGTTGTGGTGG  
 TGTGCGGCGTGCTCACTCTCTCCCTCATGAACACCCACCCACCTCGTTTCCGCAGCCCCTGC  
 GTGCTGCTCCAGAGGTGGGTGGGAGGTGAGCTGGGGGCTCCTTGGGCCCTCATCGGTGATGG  
 TCTCGTCCCATTCCACACCATTGTTTCTGTCTCCCCATCCTACTCCAAGGATGCCGGCA  
 TCACCCCTGAGGGCTCCCCCTTGGGAATGGGGTAGTGAGGCCCCAGACTTCACCCCCAGCCCA  
 CTGCTAAAATCTGTTTTCTGACAGATGGGTTTTGGGGAGTCGCCTGCTGCACTACATGAGAA  
 AGGGACTCCCATTTGCCCTTCCCTTTCTCCTACAGTCCCTTTTGTCTTGTCTGTCCTGGCTG  
 TCTGTGTGTGTGCCATTCTCTGGACTTCAGAGCCCCCTGAGCCAGTCCTCCCTTCCCAGCCT  
 CCCTTTGGGCCTCCCTAACTCCACCTAGGCTGCCAGGGACCGGAGTCAGCTGGTTCAAGGCC  
 ATCGGGAGCTCTGCCTCCAAGTCTACCCTTCCCTTCCCGGACTCCCTCCTGTCCCCTCCTTT  
 CCTCCCTCCTTCCCTTCCACTCTCCTTCCCTTTTGCTTCCCTGCCCTTTCCCCCTCCTCAGGTT  
 CTTCCCTCCTTCTCACTGGTTTTTCCACCTTCCCTCCTTCCCTTCTTCCCTGGCTCCTAGGCT  
 GTGATATATATTTTTGTATTATCTCTTTCTTCTTGTGGTGATCATCTTGAATTACTGTG  
 GGATGTAAGTTTCAAATTTTCAAATAAAGCCTTTGCAAGATAA

**FIGURE 204**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76393
><subunit 1 of 1, 243 aa, 1 stop
><MW: 26266, pI: 8.43, NX(S/T): 1
MRPQGPAASPQRLRGLLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGV
PGRDGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIA
ECTFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGGLPIEAIYLDQGSPEMN
STINIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIIEELPK
```

**Signal peptide:**

amino acids 1-30

**Transmembrane domain:**

amino acids 195-217

P  
R  
O  
T  
E  
I  
N  
S  
E  
Q  
U  
E  
N  
C  
E

**FIGURE 205**

GTTAACCAGCGCAGTCCTCCGTGCGTCCCGCCCGCCGCTGCCCTCACTCCCGGCCAGGATGG  
CATCCTGTCTGGCCCTGCGCATGGCGCTGCTGCTGGTCTCCGGGGTTCTGGCCCCTGCGGTG  
CTCACAGACGATGTTCCACAGGAGCCCGTGCCACGCTGTGGAACGAGCCGGCCGAGCTGCC  
GTCGGGAGAAGGCCCCGTGGAGAGCACCAGCCCCGGCCGGGAGCCCGTGGACACCGGTCCCC  
CAGCCCCCACCCTCGCGCCAGGACCCGAGGACAGCACC GCGCAGGAGCGGCTGGACCAGGGC  
GGCGGGTCGCTGGGGCCCGGCGCTATCGCGGCCATCGTGATCGCCGCCCTGCTGGCCACCTG  
CGTGGTGCTGGCGCTCGTGGTCGTCGCGCTGAGAAAGTTTTCTGCCTCCTGAAGCGAATAAA  
GGGGCCGCGCCCGGCCGCGGCGCGACTCGGCAAAAAAAAAAAAAA

**FIGURE 206**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76398
><subunit 1 of 1, 121 aa, 1 stop
><MW: 12073, pI: 4.11, NX(S/T): 0
MASCLALRMALLLVSGVLAPAVLTDDVPQEPVPTLWNEPAELPSGEGPVESTSPGREPVDTG
PPAPTVPAGPEDSTAQERLDQGGGSLGPGAIAAIVIAALLATCVVLALVVVALRKFSAS
```

**Important features of the protein:****Signal peptide:**

amino acids 1-19

**Transmembrane domain:**

amino acids 91-110

**Glycosaminoglycan attachment site.**

amino acids 44-47

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 116-119

**N-myristoylation site.**

amino acids 91-96

**FIGURE 207**

GGCCGTTGGTTGGTGC GCGGCTGAAGGGTGTGGCGCGAGCAGCGTCGTTGGTTGGCCGGCGG  
CGGGCCGGGACGGGGC**ATG**GCCCTGCTGCTGTGCCTGGTGTGCCTGACGGCGGCGCTGGCCCA  
CGGCTGTCTGCACTGCCACAGCAACTTCTCCAAGAAGTTCTCCTTCTACCGCCACCATGTGA  
ACTTCAAGTCCTGGTGGGTGGGCGACATCCCCGTGTCAGGGGCGCTGCTCACCGACTGGAGC  
GACGACACGATGAAGGAGCTGCACCTGGCCATCCCCGCCAAGATCACCCGGGAGAAGCTGGA  
CCAAGTGGCGACAGCAGTGTACCAGATGATGGATCAGCTGTACCAGGGGAAGATGTACTTCC  
CCGGGTATTTCCCCAACGAGCTGCGAAACATCTTCCGGGAGCAGGTGCACCTCATCCAGAAC  
GCCATCATCGAAAGGCACCTGGCACCAGGCAGCTGGGGAGGAGGGCAGCTCTCCAGGGAGGG  
ACCCAGCCTAGCACCTGAAGGATCAATGCCATCACCCCGCGGGGACCTCCCC**TAA**GTAGCCC  
CCAGAGGCGCTGGGAGTGTGGCACCGCCCTCCCCTGAAGTTTGCTCCATCTCACGCTGGGG  
GTCAACCTGGGGACCCCTTCCCTCCGGGCCATGGACACACATACATGAAAACCAGGCCGCAT  
CGACTGTCAGCACCGCTGTGGCATCTTCCAGTACGAGACCATCTCCTGCAACAACCTGCACAG  
ACTCGCACGTCGCCTGCTTTGGCTATAACTGCGAGTAGGGCTCAGGCATCACACCCACCCGT  
GCCAGGGCCCTACTGTCCCTGGGGTCCCAGGCTCTCCTTGGAGGGGGCTCCCCGCCTTCCAC  
CTGGCTGTCATCGGGTAGGGCGGGGCCGTGGGTTCAGGGGCGCACCCTTCCAAGCCTGTGT  
CCCACAGGTCCTCGGCGCAGTGGAAGTCAGCTGTCCAGGGCCTCCTGAACTACATAAATAAC  
TGGCACAAGTAAGTCCCCTCCTCAAACCAACACAGGCAGTGTGTGTATGTGAGCACCTCGTG  
GGTGAGTATGTGTGGGGCACAGGCTGGCTCCCTCAGCTCCCACGTCCTAGAGGGGCTCCCGA  
GGAGGTGGAACCTCAACCCAGCTCTGCGCAGGAGGCGGCTGCAGTCCTTTTCTCCCTCAAAG  
GTCTCCGACCCTCAGCTGGAGGCGGGCATCTTTCTAAAGGGTCCCCATAGGGTCTGGTTCC  
ACCCCATCCCAGGTCTGTGGTCAGAGCCTGGGAGGGTTCCCTACGATGGTTAGGGGTGCCCC  
ATGGAGGGGCTGACTGCCCCACATTGCCTTTCAGACAGGACACGAGCATGAGGTAAGGCCGC  
CCTGACCTGGACTTCAGGGGGAGGGGGTAAAGGGAGAGAGGAGGGGGGCTAGGGGGTCTCTCT  
AGATCAGTGGGGGCACTGCAGGTGGGGCTCTCCCTATACCTGGGACACCTGCTGGATGTCAC  
CTCTGCAACCACACCCATGTGGTGGTTTCATGAACAGACCACGTCCTCTGCCTTCTCCTGG  
CCTGGGACACACAGAGCCACCCCGGCCTTGTGAGTGACCCAGAGAAGGGAGGCCTCGGGAGA  
AGGGGTGCTCGTAAGCCAACACCAGCGTGCCGCGGCCTGCACACCCTTCGGACATCCCAGGC  
ACGAGGGTGTGCTGGATGTGGCCACACATAGGACCACACGTCCCAGCTGGGAGGAGAGGCCT  
GGGGCCCCCAGGGAGGGAGGCAGGGGGTGGGGGACATGGAGAGCTGAGGCAGCCTCGTCTCC  
CCGCAGCCTGGTATCGCCAGCCTTAAGGTGTCTGGAGCCCCCACACTTGGCCAACCTGACCT  
TGGAAGATGCTGCTGAGTGTCTCAAGCAGCACTGACAGCAGCTGGGCCTGCCCCAGGGCAAC  
GTGGGGGCGGAGACTCAGCTGGACAGCCCCGTGCTGTCACTCTGGAGCTGGGCTGCTGCTGC  
CTCAGGACCCCCTCTCCGACCCCGGACAGAGCTGAGCTGGCCAGGGCCAGGAGGGCGGGAGG  
GAGGGAATGGGGGTGGGCTGTGCGCAGCATCAGCGCCTGGGCAGGTCCGCAGAGCTGCGGGA  
TGTGATTAAAGTCCCTGATGTTTCTC

**FIGURE 208**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76399
```

```
><subunit 1 of 1, 157 aa, 1 stop
```

><MW: 17681, pI: 7.65, NX(S/T): 1

MALLLCLVCLTAALAHGCLHCHSNFSKKSFYRHHVNFKSWWVGDI PVSGALLTDWSDDTMK  
 ELHLAI PAKITREKLDQVATAVYQMMDQLYQGKMYFPGYFPNELRNIFREQVHLIQNAI IER  
 H LAPGSWGGGQLSREGPSLAPEGSMPSPRGDLP

Signal peptide:

amino acids 1-15

**FIGURE 209**

AGCAGGAGCAGGAGAGGGACA**ATG**GAAGCTGCCCCGTCCAGGTTTCATGTTCCCTCTTATTTCT  
 CCTCACGTGTGAGCTGGCTGCAGAAGTTGCTGCAGAAGTTGAGAAATCCTCAGATGGTCCTG  
 GTGCTGCCCAGGAACCCACGTGGCTCACAGATGTCCCAGCTGCCATGGAATTCATTGCTGCC  
 ACTGAGGTGGCTGTCATAGGCTTCTTCCAGGATTTAGAAATACCAGCAGTGCCCATACTCCA  
 TAGCATGGTGCAAAAATTCCCAGGCGTGTCAATTTGGGATCAGCACTGATTCTGAGGTTCTGA  
 CACACTACAACATCACTGGGAACACCATCTGCCTCTTTCGCCTGGTAGACAATGAACAACCTG  
 AATTTAGAGGACGAAGACATTGAAAGCATTGATGCCACCAAATTGAGCCGTTTCATTGAGAT  
 CAACAGCCTCCACATGGTGACAGAGTACAACCCTGTGACTGTGATTGGGTATTCAACAGCG  
 TAATTCAGATTCATCTCCTCCTGATAATGAACAAGGCCTCCCCAGAGTATGAAGAGAACATG  
 CACAGATACCAGAAGGCAGCCAAGCTCTTCCAGGGGAAGATTCTCTTTATTCTGGTGGACAG  
 TGGTATGAAAGAAAATGGGAAGGTGATATCATTTTTTCAAACATAAGGAGTCTCAACTGCCAG  
 CTTTGGCAATTTACCAGACTCTAGATGACGAGTGGGATACACTGCCCACAGCAGAAGTTTCC  
 GTAGAGCATGTGCAAACTTTTGTGATGGATTCCCTAAGTGGAATAATTGTTGAAAGAAAATCG  
 TGAATCAGAAGGAAAGACTCCAAAGGTGGAACCT**TGA**CTTCTCCTTGGAACACATATGGCC  
 AAGTATCTACTTTTATGCAAAGTAAAAAGGCACAACCTCAAATCTCAGAGACACTAAACAACAG  
 GATCACTAGGCCTGCCAACCACACACACACGACGTCACACACGACGACGCGTGCACAC  
 ACACACGCGCACACACACACACACACAGAGCTTCATTTCTGTCTTAAATCTCGTTTTCTC  
 TTCTTCCTTCTTTTAAATTTTCATATCCTCACTCCCTATCCAATTTCTTCTTATCGTGCATT  
 CATACTCTGTAAGCCCATCTGTAACACACCTAGATCAAGGCTTTAAGAGACTCACTGTGATG  
 CCTCTATGAAAGAGAGGCATTCCTAGAGAAAGATTGTTCCAATTTGTCATTTAATATCAAGT  
 TTGTATACTGCACATGACTTACACACAACATAGTTCCCTGCTCTTTTAAGGTTACCTAAGGGT  
 TGAAACTCTACCTTCTTTTATAAGCACATGTCCGTCTCTGACTCAGGATCAAAAACCAAAGG  
 ATGGTTTTAAACACCTTTGTGAAATTGTCTTTTTGCCAGAAGTTAAAGGCTGTCTCCAAGTC  
 CCTGAACTCAGCAGAAATAGACCATGTGAAAACCTCATGCTTGGTTAGCATCTCCAACCTCCC  
 TATGTAAATCAACAACCTGCATAATAATAAAAGGCAATCATGTTATA



**FIGURE 210**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76401

><subunit 1 of 1, 273 aa, 1 stop

><MW: 30480, pI: 4.60, NX(S/T): 1

MEAAPSRFMFLLFLLTCELAAEVAAEVEKSSDGPAAQEPTWLTDVPAAMEFIAATEVAVIG  
FFQDLEIPAVPILHSMVQKFPGVSEFGISTDSEVLTHYNITGNTICLFRLVDNEQLNLEDEDI  
ESIDATKLSRFIEINSLHMTVEYNPVTVIGLFNSVIQIHLLIMNKASPEYEENMHRYQKAA  
KLFQGKILFILVDSGMKENGKVISFFKLKESQLPALAIYQTLDDDEWDTLPTAEVSVEHVQNF  
CDGFLSGKLLKENRESEGKTPKVEL

**Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 143-162

## FIGURE 211

[illegible]

**FIGURE 212**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76510
><subunit 1 of 1, 269 aa, 1 stop
><MW: 29082, pI: 9.02, NX(S/T): 3
MAASAGAGAVIAAPDSRRWLWSVLAAALGLLTAGVSALEVYTPKEIFVANGTQGKLTCKFKS
TSTTGGLTSVSWSFQPEGADTTVSFFHYSQGQVYLGNYPPFKDRISWAGDLDDKSDASINIEN
MQFIHNGTYICDVKNPPDIVVQPGHIRLYVVEKENLPVFPVWVVVGIVTAVVLGLTLLISMI
LAVLYRRKNSKRDTGCGSTSESLSPVKQAPRKSPSDTEGLVKSLPSGSHQGPVIYAQLDHSG
GHHSKINKSESVVYADIRKN
```

**Signal peptide:**

amino acids 1-37

**Transmembrane domain:**

amino acids 161-183

**FIGURE 213**

GCCGGCTGTGCAGAGACGCCATGTACCGGCTCCTGTCAGCAGTGACTGCCCGGGCTGCCGCC  
 CCCGGGGGCTTGGCCTCAAGCTGCGGACGACGCGGGGTCCATCAGCGCGCCGGGCTGCCGCC  
 TCTCGGCCACGGCTGGGTGCGGGGCCTCGGGCTGGGGCTGGGGCTGGCGCTCGGGGTGAAGC  
 TGGCAGGTGGGCTGAGGGGCGCGGCCCGGCGCAGTCCCCCGCGGCCCCGACCCTGAGGCG  
 TCGCCTCTGGCCGAGCCGCCACAGGAGCAGTCCCTCGCCCCGTGGTCTCCGCAGACCCCGGC  
 GCCGCCCTGCTCCAGGTGCTTCGCCAGAGCCATCGAGAGCAGCCGCGACCTGCTGCACAGGA  
 TCAAGGATGAGGTGGGCGCACCGGGCATAGTGGTTGGAGTTTCTGTAGATGAAAAGAAGTC  
 TGGTCAGAAGGTTTAGGTTATGCTGATGTTGAGAACCGTGTACCATGTAAACCAGAGACAGT  
 TATGCGAATTGCTAGCATCAGCAAAAGTCTCACCATGGTTGCTCTTGCCAAATTGTGGGAAG  
 CAGGGAAACTGGATCTTGATATTCCAGTACAACATTATGTTCCCGAATTCCCAGAAAAAGAA  
 TATGAAGGTGAAAAGGTTTCTGTCAACAAGATTACTGATTTCCCATTTAAGTGAATTTCG  
 TCATTATGAAAAGGACATAAAAAAGGTGAAAGAAGAGAAAGCTTATAAAGCCTTGAAGATGA  
 TGAAAGAGAATGTTGCATTTGAGCAAGAAAAAGAAGGCAAAAGTAATGAAAAGAATGATTTT  
 ACTAAATTTAAACAGAGCAGGAGAATGAAGCCAAATGCCGGAATTCAAAACCTGGCAAGAA  
 AAAGAATGATTTTGAACAAGGCGAATTATATTTGAGAGAAAAGTTTGAAAATTCAATTGAAT  
 CCCTAAGATTATTTAAAAATGATCCTTTGTTCTTCAAACCTGGTAGTCAGTTTTTGTATTCA  
 ACTTTTGGCTATAACCTACTGGCAGCCATAGTAGAGAGAGCTTCAGGATGTAAATATTTGGA  
 CTATATGCAGAAAATATTCCATGACTTGGATATGCTGACGACTGTGCAGGAAGAAAACGAGC  
 CAGTGATTTACAATAGAGCAAGGTAAATGAATACCTTCTGCTGTGTCTAGCTATATCGCATC  
 TTAACACTATTTTATTAATTAAGTCAAATTTTCTTTGTTTCCATTCCAAAATCAACCTGC  
 CACATTTTGGGAGCTTTTCTACATGTCTGTTTTCTCATCTGTAAAGTGAAGGAAGTAAACA  
 TGTTTATAAAGTAAAAAAA

**FIGURE 214**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76522

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41221, pI: 8.54, NX(S/T): 0

MYRLLSAVTARAAAPGGLASSCGRRGVHQRAGLPPLGHGWVGGLGLGLGLALGVKLAGGLRG  
AAPAQSPAAPDPEASPLAEPPEQSLAPWSPQTPAPPSCRCFARAIESSRDLLHRIKDEVGA  
PGIVVGVSVDGKEVWSEGLGYADVENRVPCPETVMRIASISKSLTMVALAKLWEAGKLDLD  
IPVQHYVPEFPEKEYEGEKVSVTTRLLISHLSGIRHYEKDIKKVKEEKAYKALKMMKENVAF  
EQEKEGKSNEKNDFTKFKTEQENEAKCRNSKPGKKKNDFEQGELYLREKFENSIESLRLFKN  
DPLFFKPGSQFLYSTFGYTLLAAIVERASGCKYLDYMQKIFHDLDMLTTVQEENEPVIYNRAR

**Signal peptide:**

amino acids 1-19

**Transmembrane domain:**

amino acids 39-60

protein = 6935100

**FIGURE 215**

GTGACACTATAGAAGAGCTATGACGTCGCATGCACGCGTACGTAAGCTCGGAATTCGGCTCG  
AGGCTGGTGGGAAGAAGCCGAG**ATG**GCGGCAGCCAGCGCTGGGGCAACCCGGCTGCTCCTGC  
TCTTGCTGATGGCGGTAGCAGCGCCCAGTCGAGCCCGGGGCAGCGGCTGCCGGGCCGGGACT  
GGTGCGCGAGGGGCTGGGGCGGAAGGTCGAGAGGGCGAGGCCTGTGGCACGGTGGGGCTGCT  
GCTGGAGCACTCATTTGAGATCGATGACAGTGCCAACTTCCGGAAGCGGGGCTCACTGCTCT  
GGAACCAGCAGGATGGTACCTTGTCCCTGTCACAGCGGCAGCTCAGCGAGGAGGAGCGGGGC  
CGACTCCGGGATGTGGCAGCCCTGAATGGCCTGTACCGGGTCCGGATCCCAAGGCGACCCGG  
GGCCCTGGATGGCTGGAAGCTGGTGGCTATGTCTCCTCCTTTGTCCCTGCGTGCTCCCTGG  
TGGAGTCGCACCTGTCGGACCAGCTGACCCTGCACGTGGATGTGGCCGGCAACGTGGTGGGC  
GTGTCGGTGGTGACGCACCCCGGGGGCTGCCGGGGGCCATGAGGTGGAGGACGTGGACCTGGA  
GCTGTTCAACACCTCGGTGCAGCTGCAGCCGCCCACCACAGCCCCAGGCCCTGAGACGGCGG  
CCTTCATTGAGCGCCTGGAGATGGAACAGGCCCAGAAGGCCAAGAACCCCCAGGAGCAGAAG  
TCCTTCTTCGCCAAATACTGGATGTACATCATTTCCCGTCGTCTTCTTCATGATGTCAGG  
AGCGCCAGACACCGGGGGCCAGGGTGGGGGTGGGGGTGGGGGTGGTGGTGGGGGTAGTGGCC  
TTTGCTGTGTGCCACCCTCCCTG**TAA**GTCTATTTAAAAACATCGACGATACATTGAAATGTG  
TGAACGTTTTGAAAAGCTACAGCTTCCAGCAGCCAAAAGCAACTGTTGTTTTGGCAAGACGG  
TCCTGATGTACAAGCTTGATTGAAATTCAGTGTCACTTGATACGTTATTCAGAAACCCAAG  
GAATGGCTGTCCCCATCCTCATGTGGCTGTGTGGAGCTCAGCTGTGTTGTGTGGCAGTTTAT  
TAAACTGTCCCCCAGATCGACACGCAAAAAAAAAA

**FIGURE 216**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76529
><subunit 1 of 1, 269 aa, 1 stop
><MW: 28004, pI: 5.80, NX(S/T): 1
MAAASAGATRLLLLLLLMAVAAPSRARGSGCRAGTGARGAGAEGREGEACGTVGLLLEHSFEI
DDSANFRKRGSLLWNQQDGTLSLSQRQLSEEERGLRDVAALNGLYRVRI PR RPGALDGLEA
GGYVSSFVPACSLVESHLS DQLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDLELENTSVQ
LQPPTTAPGPETA AFIERLEMEQAQKAKNPQE QKSFFAKYWMYIIPVVLFLMMSGAPDTGGQ
GGGGGGGGGGGSGLC CVPPSL
```

**Signal peptide:**

amino acids 1-24

**Transmembrane domain:**

amino acids 226-243

**FIGURE 217**

GGAGCGCTGCTGGAACCCGAGCCGGAGCCGGAGCCACAGCGGGGAGGGTGGCCTGGCGGCCT  
 GGAGCCGGACGTGTCCGGGGCGTCCCCGAGACCGGGGAGCAGGTCGTCCGGGGGGCCACC  
ATGCTGGTGACTGCCTACCTTGCTTTTGTAGGCCTCCTGGCCTCCTGCCTGGGGCTGGAAC  
 GTCAAGATGCCGGGCTAAACCCCTGGAAGGGCCTGCAGCAATCCCTCCTTCCTTCGGTTTC  
 AACTGGACTTCTATCAGGTCTACTTCCTGGCCCTGGCAGCTGATTGGCTTCAGGCCCCCTAC  
 CTCTATAAACTCTACCAGCATTACTACTTCCTGGAAGGTCAAATTGCCATCCTCTATGTCTG  
 TGGCCTTGCCCTCTACAGTCCTCTTTGGCCTAGTGGCCTCCTCCCTTGTGGATTGGCTGGGTC  
 GCAAGAATTCTTGTGTCTCTTCTCCCTGACTTACTACTATGCTGCTTAACCAAACCTCTCT  
 CAAGACTACTTTGTGCTGCTAGTGGGGCGAGCACTTGGTGGGCTGTCCACAGCCCTGCTCTT  
 CTCAGCCTTCGAGGCCTGGTATATCCATGAGCACGTGGAACGGCATGACTTCCCTGCTGAGT  
 GGATCCCAGCTACCTTTGCTCGAGCTGCCTTCTGGAACCATGTGCTGGCTGTAGTGGCAGGT  
 GTGGCAGCTGAGGCTGTAGCCAGCTGGATAGGGCTGGGGCCTGTAGCGCCCTTTGTGGCTGC  
 CATCCCTCTCCTGGCTCTGGCAGGGGGCCTTGGCCCTTCGAAACTGGGGGGAGAACTATGACC  
 GGCAGCGTGCCTTCTCAAGGACCTGTGCTGGAGGCCTGCGCTGCCTCCTGTCTGGACCGCCGC  
 GTGCTGCTGCTGGGCACCATAACAAGCTCTATTTGAGAGTGTCTATCTTCATCTTTGTCTTCCT  
 CTGGACACCTGTGCTGGACCCACACGGGGGCCCTCTGGGCATTATCTTCTCCAGCTTCATGG  
 CAGCCAGCCTGCTTGGCTCTTCCCTGTACCGTATCGCCACCTCCAAGAGGTACCACCTTCAG  
 CCCATGCACCTGCTGTCCCTTGCTGTGCTCATCGTCTTCTCTCTCTTCATGTTGACTTT  
 CTCTACCAGCCCAGGCCAGGAGAGTCCGGTGGAGTCCTTCATAGCCTTTCTACTTATTGAGT  
 TGGCTTGTGGATTATACTTTCCAGCATGAGCTTCCTACGGAGAAAGGTGATCCCTGAGACA  
 GAGCAGGCTGGTGTACTCAACTGGTTCCGGGTACCTCTGCACTCACTGGCTTGCCTAGGGCT  
 CCTTGTCCTCCATGACAGTGATCGAAAAACAGGCACTCGGAATATGTTACAGATTTGCTCTG  
 CTGTCATGGTGATGGCTCTGCTGGCAGTGGTGGGACTCTTCACCGTGGTAAGGCATGATGCT  
 GAGCTGCGGGTACCTTCACCTACTGAGGAGCCCTATGCCCTGAGCTGTAACCCCACTCCAG  
 GACAAGATAGCTGGGACAGACTCTTGAATTCCAGCTATCCGGGATTGTACAGATCTCTCTGT  
 GACTGACTTTGTGACTGTCCTGTGGTTTCTCCTGCCATTGCTTTGTGTTTGGGAGGACATGA  
 TGGGGGTGATGGACTGGAAAGAAGGTGCCAAAAGTTCCCTCTGTGTTACTCCCATTTAGAAA  
 ATAAACACTTTTAAATGATCAAAAAAAAAAAAA



**FIGURE 218**

MLVTAYLAFVGLLASCLGLELSRCRAKPPGRACSNPSFLRFQLDFYQVYFLALAADWLQAPY  
LYKLYQHYYFLEGQIAILYVCGLASTVLFGLVASSLVDWLGRKNSCVLFSLTYSLCCLTKLS  
QDYFVLLVGRALGGLSTALLFSAFEAWYIHEHVERHDFPAEWIPATFARAAFWNHVLAVVAG  
VAAEAVASWIGLGPVAPFVAAIPLLALAGALALRNWGENYDRQRAFSRTCAGGLRCLLSDRR  
VLLLGTIQALFESVIFIFVFLWTPVLDPHGAPLGIIFFSSFMAASLLGSSLYRIATSKRYHLQ  
PMHLLSLAVLIVVFSLFMLTFSTSPGQESPVESFIAFLLLIELACGLYFPSMSFLRRKVIPET  
EQAGVLNWFVRVPLHSLACLGLLVLHDSRKTGTRNMFSCSAVMVMALLAVVGLFTVVVRHDA  
ELRVPSPTTEEPYAPEL

**Signal peptide:**

amino acids 1-18

**Transmembrane domain:**

amino acids 41-55, 75-94, 127-143, 191-213, 249-270, 278-299,  
314-330, 343-359, 379-394, 410-430

**FIGURE 219**

GCGACGCGCGGCGGGGCGGCGAGAGGAAACGCGGCGCCGGGCGGGCCCGGCCCTGGAGATG  
 GTCCCCGGCGCCGCGGGCTGGTGTGTCTCGTGCTCTGGCTCCCCGCGTGCGTCGCGGCCCA  
 CGGCTTCCGTATCCATGATTATTTGTACTTTCAAGTGCTGAGTCCTGGGGACATTCGATACA  
 TCTTCACAGCCACACCTGCCAAGGACTTTGGTGGTATCTTTCACACAAGGTATGAGCAGATT  
 CACCTTGTCCCCGCTGAACCTCCAGAGGCCTGCGGGGAACTCAGCAACGGTTTCTTCATCCA  
 GGACCAGATTGCTCTGGTGGAGAGGGGGGGCTGCTCCTTCCTCTCCAAGACTCGGGTGGTCC  
 AGGAGCACGGCGGGCGGGCGGTGATCATCTCTGACAACGCAGTTGACAATGACAGCTTCTAC  
 GTGGAGATGATCCAGGACAGTACCCAGCGCACAGCTGACATCCCCGCCCTCTTCCTGCTCGG  
 CCGAGACGGCTACATGATCCGCCGCTCTCTGGAACAGCATGGGCTGCCATGGGCCATCATTT  
 CCATCCCAGTCAATGTCACCAGCATCCCCACCTTTGAGCTGCTGCAACCGCCCTGGACCTTC  
 TGGTAGAAAGAGTTTGTCCACATTCCAGCCATAAGTGACTCTGAGCTGGGAAGGGGAAACCC  
 AGGAATTTTGCTACTTGAATTTGGAGATAGCATCTGGGGACAAGTGGAGCCAGGTAGAGGA  
 AAAGGGTTTGGGCGTTGCTAGGCTGAAAGGGAAGCCACACCACTGGCCTTCCCTTCCCCAGG  
 GCCCCAAGGGTGTCTCATGCTACAAGAAGAGGCAAGAGACAGGCCCCAGGGCTTCTGGCTA  
 GAACCCGAAACAAAAGGAGCTGAAGGCAGGTGGCCTGAGAGCCATCTGTGACCTGTCACACT  
 CACCTGGCTCCAGCCTCCCCTACCCAGGGTCTCTGCACAGTGACCTTCACAGCAGTTGTTGG  
 AGTGGTTTAAAGAGCTGGTGTGTTGGGGACTCAATAAACCCCTCACTGACTTTTTTAGCAATAAA  
 GCTTCTCATCAGGGTTGCAAAAAAAAAAAAAAAAAAAAAAAAAA



**FIGURE 221**

TCTGCCTCCACTGCTCTGTGCTGGGATC**ATG**GAACTTGCACTGCTGTGTGGGCTGGTGGTGA  
TGGCTGGTGTGATTCCAATCCAGGGCGGGATCCTGAACCTGAACAAGATGGTCAAGCAAGTG  
ACTGGGAAAATGCCCATCCTCTCCTACTGGCCCTACGGCTGTCACTGCGGACTAGGTGGCAG  
AGGCCAACCCAAAGATGCCACGGACTGGTGCTGCCAGACCCATGACTGCTGCTATGACCACC  
TGAAGACCCAGGGGTGCGGCATCTACAAGGACAACAACAAAAGCAGCATACATTGTATGGAT  
TTATCTCAACGCTATTGTTTAATGGCTGTGTTTAATGTGATCTATCTGGAAAATGAGGACTC  
CGAA**TAAA**AAGCTATTACTAWTTNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAA

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250

**FIGURE 222**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76538
><subunit 1 of 1, 116 aa, 1 stop
><MW: 12910, pI: 6.41, NX(S/T): 1
MELALLCGLVVMAGVPIPIQGILNLNKMVKQVTGKMPILSYWPYGCHCGLGGRGQPKDATDW
CCQTHDCCYDHLKTQGCgiYKDNNKSSIHCMdLSQRYCLMAVFNViiYLENEdSE
```

**Important features of the protein:****Signal peptide:**

amino acids 1-17

**Transmembrane domain:**

amino acids 1-24

**N-glycosylation site.**

amino acids 86-89

**N-myristoylation sites.**

amino acids 20-25, 45-50

**Phospholipase A2 histidine active site.**

amino acids 63-70

**FIGURE 223**

CTCGCTTCTTCCTTCTGGATGGGGGGCCAGGGGGCCAGGAGAGTATAAAGGCGATGTGGAG  
GGTGCCCGGCACAACCAGACGCCAGTCACAGGCGAGAGCCCTGGGATGCACCGGCCAGAGG  
CCATGCTGCTGCTGCTCACGCTTGCCCTCCTGGGGGGCCCCACCTGGGCAGGGAAGATGTAT  
GGCCCTGGAGGAGGCAAGTATTTTCAGCACCACTGAAGACTACGACCATGAAATCACAGGGCT  
GCGGGTGTCTGTAGGTCTTCTCCTGGTGAAAAGTGTCCAGGTGAACTTGGAGACTCCTGGG  
ACGTGAAACTGGGAGCCTTAGGTGGGAATACCCAGGAAGTCACCCTGCAGCCAGGCGAATAC  
ATCACAAAAGTCTTTGTGCGCTTCCAAGCTTTCCTCCGGGGTATGGTCATGTACACCAGCAA  
GGACCGCTATTTCTATTTTGGGAAGCTTGATGGCCAGATCTCCTCTGCCTACCCCAGCCAAG  
AGGGGCAGGTGCTGGTGGGCATCTATGGCCAGTATCAACTCCTTGGCATCAAGAGCATTGGC  
TTTGAATGGAATTATCCACTAGAGGAGCCGACCACTGAGCCACCAGTTAATCTCACATACTC  
AGCAAACCTACCCGTGGGTCGCTAGGGTGGGGTATGGGGCCATCCGAGCTGAGGCCATCTGT  
GTGGTGGTGGCTGATGGTACTGGAGTAACTGAGTCGGGACGCTGAATCTGAATCCACCAATA  
AATAAAGCTTCTGCAGAAAA

**FIGURE 224**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76541
><subunit 1 of 1, 178 aa, 1 stop
><MW: 19600, pI: 5.89, NX(S/T): 1
MHRPEAMLLLLTLALLGGPTWAGKMYGPGGGKYFSTTEDYDHEITGLRVSVGLLLKSVQVK
LGDSWDVKLGALGGNTQEVTLQPGEYITKVVFVAFQAFLRGMVMYTSKDRYFYFGKLDGQISS
AYPSQEGQVLVGIYGQYQLLGIKSIGFEWNYPLEEPTTEPPVNLTYSANSPVGR
```

**Signal peptide:**

amino acids 1-22

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250

## FIGURE 225

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAAGCTCTGTGGCT  
GAACTGGGTGCTCATCACGGGAAGTCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAG  
CCCCAAATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTT  
TTTTTTTTTAACCGCCCCCTCCCCACCCCCCAAAAAAAGTGTAAAGATGCAAAAACGTAATAT  
CCATGAAGATCCTATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATT  
TATTTGTTCTTGAGTGTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCT  
CCCAAGGGGTCCAATTTTTCTTCCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTG  
ACAGGGGCTGTCATGCAACTGGCCCCTAAGCCAAAGCAAAGACCTAAGGACGACCTTTGAA  
CAATACAAAGGATGGGTTTCAATGTAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTT  
ATAGCCCCCACTGTCTTACTGACAATGCTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTG  
TAGGTGTGAAGGCAAATGGTATATTGTGAATCTCAGAAATTACAGGAGATACCCTCAAGTA  
TATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCTATAACAGCCTTCAAAAAGTATAAT  
CAATTTAAAGGGCTCAACCAGCTCACCTGGCTATACCTTGACCATAACCATATCAGCAATAT  
TGACGAAAATGCTTTTAATGGAATACGCAGACTCAAAGAGCTGATTCTTAGTTCCAATAGAA  
TCTCCTATTTTCTTAACAATACCTTCAGACCTGTGACAAATTTACGGAACTTGGATCTGTCC  
TATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTGCGGAAGCTGCTGAGTTT  
ACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCAAGACTGCCGCAACC  
TGGAACTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATGTCTTTGCTGGC  
ATGATCAGACTCAAAGAAGTTTACCTGGAGCACAATCAATTTTCCAAGCTCAACCTGGCCCT  
TTTTCCAAGGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGGAATAAAATCAGTGTCTATAG  
GACAGACCATGTCTTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGATC  
GAAGCTTTTCAGTGGACCCAGTGTTTTCCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGA  
TTCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACA  
TCAGTCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTTGTAAGTGGCTG  
AAAAGTTTTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGG  
AGTAAATGTGATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAGTACTACAGAGAGGT  
TTGATCTGGCCAGGGCTCTCCCAAAGCCGACGTTTAAAGCCCAAGCTCCCCAGGCCGAAGCAT  
GAGAGCAAACCCCTTTGCCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGC  
TGACGCCGAGCACATCTTTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCG  
TGCTCGTCATCCTGCTGGTTATCTACGTGTCATGGAAGCGGTACCCTGCGAGCATGAAGCAG  
CTGCAGCAGCGCTCCCTCATGCGAAGGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAAT  
GACTCCAGCACCCAGGAATTTTATGTAGATTATAAACCACCAACACGGAGACCAGCGAGA  
TGCTGCTGAATGGGACGGGACCCTGCACCTATAACAAATCGGGCTCCAGGGAGTGTGAGGTA  
**TGA**ACCATTGTGATAAAAAGAGCTCTTAAAGTGGGAAATAAGTGGTGCTTTATTGAACTC  
TGGTGACTATCAAGGGAACGCGATGCCCCCCTCCCCTTCCCTCTCCCTCTCACTTTGGTG  
CAAGATCCTTCCTTGTCCGTTTTAGTGCATTCTATAATACTGGTCATTTTCTCTCATACATA  
ATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTTGAAGTCCGGTTTAATATAA  
TACCTATTGTATAAGACCTTTACTGATTCCATTAATGTGCGATTTGTTTTTAAGATAAACT  
TCTTTCATAGGTAAAAA



**FIGURE 226**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77301
><subunit 1 of 1, 513 aa, 1 stop
><MW: 58266, pI: 9.84, NX(S/T): 4
MGFNVIRLLSGSAVALVIAPTIVLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAG
CLGLSLRYNSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYF
LNNTFRPVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRIFQDCRNLELL
DLGYNRIRSLARNVFAGMIRLKEHLEHNQFSKLNALFPRLVSLQONLYLQWNKISVIGQTM
SWTWSSLQRLDLSGNEIEAFSGPSVFQCVPNLQRLNLDLSNKLTFIGQEILDSWISLNDISLA
GNIWECSRNICSLVNWLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLA
RALPKPTFKPKLPRPKHESKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVI
LLVIYVSWKRYPASMKQLQQRSLMRRHRKKKRQSLKQMTPTSTQEFYVDYKPTNTTETSEMLLN
GTGPCTYNKSGSRECEV
```

**Important features of the protein:****Signal peptide:**

amino acids 1-33

**Transmembrane domain:**

amino acids 420-442

**N-glycosylation sites.**

amino acids 126-129, 357-360, 496-499, 504-507

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 465-468

**Tyrosine kinase phosphorylation site.**

amino acids 136-142

**N-myristoylation sites.**

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

**FIGURE 227**

AGTTCTGAGAAAGAAGGAAATAAACACAGGCACCAAACCACTATCCTAAGTTGACTGTCCTT  
 TAAAT**ATG**TCAGATCCAGACTTTTCAGTGTACCTCAGCGATCTCAACGATAGGGATCTTG  
 TGTTTGCCGCTATTCCAGTTGGTGCTCTCGGACCTACCATGCGAAGAAGATGAAATGTGTGT  
 AAATTATAATGACCAACACCCTAATGGCTGGTATATCTGGATCCTCCTGCTGCTGGTTTTGG  
 TGGCAGCTCTTCTCTGTGGAGCTGTGGTCCTCTGCCTCCAGTGCTGGCTGAGGAGACCCCGA  
 ATTGATTCTCACAGGCGCACCATGGCAGTTTTTGGCTGTTGGAGACTTGGACTCTATTTATGG  
 GACAGAAGCAGCTGTGAGTCCAACTGTTGGAATTCACCTTCAAACCTCAAACCCCTGACCTAT  
 ATCCTGTTCTGCTCCATGTTTTGGCCCTTTAGGCTCCCCACCTCCATATGAAGAAATTGTA  
 AAAACAACCT**TGA**TTTTAGGTGTGGATTATCAATTTAAAGTATTAACGACATCTGTAATTCCA  
 AAACATCAAATTTAGGAATAGTTATTTTCAGTTGTTGGAAATGTCCAGAGATCTATTCATATA  
 GTCTGAGGAAGGACAATTCGACAAAAGAATGGATGTTGGAAAAAATTTTGGTCATGGAGATG  
 TTTAAATAGTAAAGTAGCAGGCTTTTGTATGTGTCACTGCTGTATCATACTTTTATGCTACAC  
 AACCAAATTAATGCTTCTCCACTAGTATCCAAACAGGCAACAATTAGGTGCTGGAAGTAGTT  
 TCCATCACATTTAGGACTCCACTGCAGTATACAGCACACCATTTTCTGCTTTAAACTCTTTC  
 CTAGCATGGGGTCCATAAAAAATTATTATAATTTAACAATAGCCCAAGCCGAGAATCCAACAT  
 GTCCAGAACCAGAACCAGAAAGATAGTATTTGAATGAAGGTGAGGGGAGAGAGTAGGAAAAA  
 GAAAAGTTTGGAGTTGAAGGGTAAAGGATAAATGAAGAGGAAAAGGAAAAGATTACAAGTCT  
 CAGCAAAAACAAGAGGTTTTATGCCCCAACCTGAAGAGGAAGAAATTGTAGATAGAAGGTGA  
 AGGAGATTGCTGAAGATATAGAGCACATATAATGCCAACACGGGGAGAAAAGAAAATTTCCC  
 CTTTTACAGTAATGAATGTGGCCTCCATAGTCCATAGTGTTTCTCTGGAGCCTCAGGGCTTG  
 GCATTTATTGCAGCATCATGCTAAGAACCTTCGGCATAGGTATCTGTTCCCATGAGGACTGC  
 AGAAGTAGCAATGAGACATCTTCAAGTGGCATTTTGGCAGTGGCCATCAGCAGGGGGACAGA  
 CAAAACATCCATCACAGATGACATATGATCTTCAGCTGACAAATTTGTTGAACAAAACAAT  
 AAACATCAATAGATATCTAAAAA

**FIGURE 228**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77303

><subunit 1 of 1, 146 aa, 1 stop

><MW: 16116, pI: 4.99, NX(S/T): 0

MSRSRLFSVTSAISTIGILCLPLFQLVLSDLPCEEDEMCVNYNDQHPNGWYIWILLLLVLVA  
ALLCGAVVLCLQCWLRRPRIDSHRRTMAVFAVGDLDSIYGTEAAVSPTVGIHLQTQTPDLYP  
VPAPCFGPLGSPPPYEEIVKTT

**Signal peptide:**

amino acids 1-29

**Transmembrane domain:**

amino acids 52-70

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100



**FIGURE 230**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77648
><subunit 1 of 1, 140 aa, 1 stop
><MW: 15668, pI: 10.14, NX(S/T): 5
MFFTISRKNMSQKLSLLLLVFGLIWGLMLLHYTFQQPRHQSSVKLREQILDLSKRYVKALAE
ENKNTVDVENGASMAGYADLKRTIAVLLDDILQRLVKLENKVDYIVVNGSAANTTNGTSGNL
VPVTTNKRTNVSGSIR
```

**Important features of the protein:**

**Signal peptide:**

amino acids 1-26

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 231**

CGCGGCCGGGCGCGCGGGGTGAGCGTGCCGAGGCGGCTGTGGCGCAGGCTTCCAGCCCCAC  
**CATG**CCCGTGGCCCCTGCTGCTGCTGCTGGCCGTGAGTGGGGCCAGACAACCCGGCCATGCT  
TCCCCGGGTGCCAATGCGAGGTGGAGACCTTCGGCCTTTTCGACAGCTTCAGCCTGACTCGG  
GTGGATTGTAGCGGCCTGGGCCCCACATCATGCCGGTGCCCATCCCTCTGGACACAGCCCA  
CTTGGACCTGTCCTCCAACCGGCTGGAGATGGTGAATGAGTCGGTGTTGGCGGGGCGGGGCT  
ACACGACGTTGGCTGGCCTGGATCTCAGCCACAACCTGCTCACCAGCATCTCACCCTGCTGCC  
TTCTCCCGCCTTCGCTACCTGGAGTCGCTTGACCTCAGCCACAATGGCCTGACAGCCCTGCC  
AGCCGAGAGCTTCACCAGCTCACCCTGAGCGACGTGAACCTTAGCCACAACCAGCTCCGGG  
AGGTCTCAGTGTCTGCCTTCACGACGCACAGTCAGGGCGGGGCACTACACGTGGACCTCTCC  
CACAACCTCATTACCGCCTCGTGCCCCACCCACGAGGGCGGCCTGCCTGCGCCACCAT  
TCAGAGCCTGAACCTGGCCTGGAACCGGCTCCATGCCGTGCCAACCTCCGAGACTTGCCCC  
TGCGCTACCTGAGCCTGGATGGGAACCTCTAGCTGTCAATTGGTCCGGGTGCCTTCGCGGGG  
CTGGGAGGCCTTACACACCTGTCTCTGGCCAGCCTGCAGAGGCTCCCTGAGCTGGCGCCAG  
TGGCTTCCGTGAGCTACCGGGCCTGCAGGTCTGGACCTGTGCGGCAACCCCAAGCTTAAC  
GGGAGGAGCTGAGGTGTTTTAGGCCTGAGCTCCCTGCAGGAGCTGGACCTTTCGGGCACC  
AACCTGGTGCCCTGCCTGAGGCGCTGCTCCTCCACCTCCCGGCACTGCAGAGCGTCAGCGT  
GGGCCAGGATGTGCGGTGCCGGCGCCTGGTGCGGGAGGGCACCTACCCCGGAGGCCTGGCT  
CCAGCCCCAAGGTGCCCTGCACTGCGTAGACACCCGGGAATCTGCTGCCAGGGGCCCCACC  
ATCTT**TGA**CAAATGGTGTGGCCAGGGCCACATAACAGACTGCTGTCTCTGGGCTGCCTCAG  
GTCCCGAGTAACTTATGTTCAATGTGCCAACACCAGTGGGGAGCCCGCAGGCCTATGTGGCA  
GCGTCACCACAGGAGTTGTGGGCCTAGGAGAGGCTTTGGACCTGGGAGCCACACCTAGGAGC  
AAAGTCTCACCCCTTTGTCTACGTTGCTTCCCCAAACCATGAGCAGAGGGACTTCGATGCCA  
AACCAGACTCGGGTCCCCTCCTGCTTCCCTTCCCCACTTATCCCCCAAGTGCCTTCCCTCAT  
GCCTGGGCGGCCTGACCCGCAATGGGCAGAGGGTGGGTGGGACCCCTGCTGCAGGGCAGA  
GTTCAGGTCCACTGGGCTGAGTGTCCCCTTGGGCCCATGGCCAGTCACTCAGGGGCGAGTT  
TCTTTTCTAACATAGCCCTTTCTTTGCCATGAGGCCATGAGGCCCGCTTCATCCTTTTCTAT  
TTCCCTAGAACCTTAATGGTAGAAGGAATTGCAAAGAATCAAGTCCACCCTTCTCATGTGAC  
AGATGGGGAACTGAGGCCTTGAGAAGGAAAAAGGCTAATCTAAGTTCCTGCGGGCAGTGGC  
ATGACTGGAGCACAGCCTCCTGCCTCCCAGCCCGGACCCAATGCACTTTCTTGTCTCCTCTA  
ATAAGCCCCACCCTCCCCGCTGGGCTCCCCCTTGCTGCCCTTGCTGTTCCCCATTAGCACA  
GGAGTAGCAGCAGCAGGACAGGCAAGAGCCTCACAAGTGGGACTCTGGGCCTCTGACCAGCT  
GTGCGGCATGGGCTAAGTCACTCTGCCCTTCGGAGCCTCTGGAAGCTTAGGGCACATTGGTT  
CCAGCCTAGCCAGTTTCTCACCTGGGTTGGGGTCCCCCAGCATCCAGACTGGAAACCTACC  
CATTTTCCCCTGAGCATCCTCTAGATGCTGCCCCAAGGAGTTGCTGCAGTTCTGGAGCCTCA  
TCTGGCTGGGATCTCCAAGGGGCCTCCTGGATTAGTCCCCACTGGCCCTGAGCACGACAGC  
CCTTCTTACCCTCCCAGGAATGCCGTGAAAGGAGACAAGGTCTGCCCGACCCATGTCTATGC  
TCTACCCCCAGGGCAGCATCTCAGCTTCCGAACCTGGGCTGTTTCCTTAGTCTTCATTTTA  
TAAAAGTTGTTGCCTTTTTAACGGAGTGTCACCTTCAACCGGCCTCCCCTACCCCTGCTGGC  
CGGGGATGGAGACATGTCATTTGTAAAAGCAGAAAAAGGTTGCATTTGTTCACTTTTGTAA  
ATTGTCTGGGCTGTGTTGGGGTGTGGGGGAAGCTGGGCATCAGTGGCCACATGGGCATC  
AGGGGCTGGCCCCACAGAGACCCACAGGGCAGTGAGCTCTGTCTTCCCCACCTGCCTAGC  
CCATCATCTATCTAACCGGTCTTGATTTAATAAACACTATAAAAGGTTTAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 232**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77652

><subunit 1 of 1, 353 aa, 1 stop

><MW: 37847, pI: 6.80, NX(S/T): 2

MPWPLLLLLLAVSGAQTTTRPCFPGCQCEVETFGLFDSFSLTRVDCSGLGPHIMPVPIPLDTAH  
LDLSSNRLEMVNESVLAGPGYTTLAGLDLSHNLLTSISPTAFSRLRYLESLDLSHNGLTALP  
AESFTSSPLSDVNLSHNQLREVSVSAFTTHSQGRALHVDLSHNLIHRLVPHPTRAGLPAPTI  
QSLNLAWNRLHAVPNLRDLPLRYLSLDGNPLAVIGPGAFAAGLGGGLTHLSLASLQRLPELAPS  
GFRELPGQLQVLDLSGNPKLNWAGAEVFSGLSLQELDLSGTNLVPLPEALLLHLPALQSVSV  
GQDVRCRRLVREGTYPRRPGSSPKVPLHCVDTRESAARGPTIL

**Signal peptide:**

amino acids 1-16

**Transmembrane domains:**

amino acids 215-232, 287-304

**FIGURE 233**

GATGGCGCAGCCACAGCTTCTGTGAGATTTCGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGG  
 GGACCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCA  
 AAACAAGTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCT  
 GTTCCAGGCCTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGC  
 CACGACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCAC  
 GGGAGGCTTGGCAGTTTTTCTTACTCCTGTGGTCTCCAGATTTTCAGGCCTAAGATGAAAGCC  
 TCTAGTCTTGCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGG  
 ACTGAAGACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATG  
 GATTTTCTGAGATACGGGGCAGTGTGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTA  
 AGGAGGACTGAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTT  
 GCTAAGACTCTATCTGGACAGGGTATTTAAAAACTACCAGACCCCTGACCATTATACTCTCC  
 GGAAGATCAGCAGCCTCGCCAATTCCTTTCTTACCATCAAGAAGGACCTCCGGCTCTCTCAT  
 GCCCACATGACATGCCATTGTGGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCA  
 CTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGGAACTAGACATTCTTC  
 TGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTGCTGCTAAGAATATTCGAGGT  
 CAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAAACCACCATCTCTTTACT  
 GTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCATGAT  
 TGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTTTTGTAATATCTT  
 TCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTTTGCTATTTAATGTATTT  
 ATTTTTTTTACTTGACATGAACTTTAAAAAAATTCACAGATTATATTTATAACCTGACTAG  
 AGCAGGTGATGTATTTTTTATACAGTAAAAAAAAAACCTTGTAATTCTAGAAGAGTGGCT  
 AGGGGGGTATTTCATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGA  
 TATTTGAAATTGAACCAATGACTACTTAGGATGGGTGTGGAATAAGTTTTGATGTGGAATT  
 GCACATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATATTGTGTAT  
 CTTCCAGCCAGGAATCCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATA  
 CCAAAAAAAAAAAAAAAAAAAAAA



**FIGURE 234**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA83500
><subunit 1 of 1, 261 aa, 1 stop
><MW: 29667, pI:.8.76, NX(S/T): 0
MRQFPKTSFDISPEMSFSIYSLQVPAVPG LTCWALTAE PGWGQNGATT CATNSHSDSEL RP
EIFSSREAWQFFLLLWSPDFRPKM KASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNL
QEIRNGFSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTP
DHYTLRKISSLAN SFLT IKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALG
ELDILLQWMEETE
```

**Important features of the protein:****Signal peptide:**

amino acids 1-42

**cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 192-195, 225-228

**N-myristoylation sites.**

amino acids 42-47, 46-51, 136-141

**FIGURE 235**

CCGTTATCGTCTTGCGCTACTGCTGA**ATG**TCCGTCCCGGAGGAGGAGGAGAGGCTTTTGCCG  
CTGACCCAGAGATGGCCCCGAGCGAGCAAATTCCCTACTGTCCGGCTGCGCGGCTACCGTGGC  
CGAGCTAGCAACCTTTCCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAAGCAG  
CTCTTGCTCGGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACA  
GCCCTAGGGATCATTTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCAT  
TTACAGACACGTAGTGTATTCTGGAGGTGGAATGGTCACATATGAACATCTCCGAGAGGTTG  
TGTTTGGCAAAAGTGAAGATGAGCATTATCCCCTTTGGAAATCAGTCATTGGAGGGATGATG  
GCTGGTGTTATTGGCCAGTTTTTTAGCCAATCCAACTGACCTAGTGAAGGTTTCAGATGCAAAT  
GGAAGGAAAAAGGAAACTGGAAGGAAAACCATTCGATTTCGTGGTGTACATCATGCATTTG  
CAAAAATCTTAGCTGAAGGAGGAATACGAGGGCTTTGGGCAGGCTGGGTACCCAATATACAA  
AGAGCAGCACTGGTGAATATGGGAGATTTAACCACCTTATGATACAGTGAAACACTACTTGGT  
ATTGAATACACCACTTGAGGACAATATCATGACTCACGGTTTATCAAGTTTATGTTCTGGAC  
TGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCATCAAAGCAGAATAATGAATCAACCA  
CGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCGACTGACTGCTTGATTCAGGCTGT  
TCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTTACCATCTTGGCTGAGAATGACCC  
CTTGGTCAATGGTGTCTGGCTTACTTATGAAAAAATCAGAGAGATGAGTGGAGTCAGTCCA  
TTT**TAA**

**FIGURE 236**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77568
><subunit 1 of 1, 323 aa, 1 stop
><MW: 36064, pI: 9.33, NX(S/T): 1
MSVPEEEEERLLPLTQRWPRASKFLLSGCAATVAELATFFPLDLTKTRLQMQGEEAALARLGDGA
RESAPYRGMVRTALGIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLREVVFVKSEDEH
YPLWKSIVIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGKPLRFRGVHHAFKILAEGGI
RGLWAGWVPNIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTP
ADVIKSRIMNQPRDKQGRGLLYKSSTDCLIQAVQGEFMSLYKGFLPSWLRMTPWMSVFWLT
YEKIREMSGVSPF
```

**Transmembrane domains:**

amino acids 25-38, 130-147, 233-248

**FIGURE 237**

CGGACGCGTGGGCGCGGGACGCCGGCAGGGTTGTGGCGCAGCAGTCTCCTTCCTGCGCGCGC  
 GCCTGAAGTCGGCGTGGGCGTTTGAGGAAGCTGGGATACAGCATTTAATGAAAAATTTATGC  
 TTAAGAAGTAAAA**ATG**GCAGGCTTCCTAGATAATTTTCGTTGGCCAGAATGTGAATGTATTG  
 ACTGGAGTGAGAGAAGAAATGCTGTGGCATCTGTTGTGCGAGGTATATTGTTTTTTACAGGC  
 TGGTGGATAATGATTGATGCAGCTGTGGTGTATCCTAAGCCAGAACAGTTGAACCATGCCTT  
 TCACACATGTGGTGTATTTTCCACATTGGCTTTCTTCATGATAAATGCTGTATCCAATGCTC  
 AGGTGAGAGGTGATAGCTATGAAAGCGGCTGTTTAGGAAGAACAGGTGCTCGAGTTTGGCTT  
 TTCATTGGTTTTCATGTTGATGTTTGGGTCACCTATTGCTTCCATGTGGATTCTTTTTTGGTGC  
 ATATGTTACCCAAAATACTGATGTTTATCCGGGACTAGCTGTGTTTTTTTCAAAATGCACTTA  
 TATTTTTTTAGCACTCTGATCTACAAATTTGGAAGAACCGAAGAGCTATGGACCT**GA**GATCAC  
 TTCTTAAGTCACATTTTCTTTTTGTTATATTCTGTTTGTAGATAGGTTTTTTATCTCTCAGT  
 ACACATTGCCAAATGGAGTAGATTGTACATTAAATGTTTTGTTTCTTTACATTTTTTATGTTT  
 TGAGTTTTGAAATAGTTTTATGAAATTTCTTTATTTTTTCATTGCATAGACTGTTAATATGTA  
 TATAATACAAGACTATATGAATTGGATAATGAGTATCAGTTTTTTTATTCTTGAGATTTAGAA  
 CTTGATCTACTCCCTGAGCCAGGGTTACATCATCTTGTCAATTTTAGAAGTAACCACTCTTGT  
 CTCTCTGGCTGGGCACGGTGGCTCATGCCTGTAATCCCAGCACCTTGGGAGGCCGAGGCGGG  
 CCGATTGCTTGAGGTCAAGTGTGTTGAGACCAGCCTGGCCAACATGGCGAAACCCCATCTACT  
 AAAAATACAAAAATTAGCCAGGCATGGTGGTGGGTGCCTGTAATCCCAGCTACCTGGGAGGC  
 TGAGGCAGGAGAATCGCTTGAACCCGGGGGGCAGAGGTTGCAGTGAGCTGAGTTTGCGCCAC  
 TGCACCTCTAGCCTGGGGGAGAAAGTGAAACTCCCTCTCAAAAAAAGACCACTCTCAGTATC  
 TCTGATTTCTGAAGATGTACAAAAAATATAGCTTCATATATCTGGAATGAGCACTGAGCCA  
 TAAAAGGTTTTTCAAGCAAGTTGTAACCTATTTTGGCCTAAAAATGAGGTTTTTTTTGGTAAAGA  
 AAAAATATTTGTTCTTATGTATTGAAGAAGTGTACTTTTATATAATGATTTTTTAAATGCCC  
 AAAGGACTAGTTTGAAAGCTTCTTTTAAAAAGAATTCCTCTAATATGACTTTATGTGAGAA

## **FIGURE 238**

MAGFLDNFRWPECECIDWSERRNAVASVVAGILFFTGWIMIDAADVYPKPEQLNHAFHTCG  
VFSTLAFFMINAVSNAQVRGDSYESGCLGRTGARVWLFIGFMLMFGSLIASMWILFGAYVTQ  
NTDVYPGLAVFFQNALIFFSTLIYKFG RTEELWT

**Important features:**

**Signal peptide:**

amino acids 1-44

**Transmembrane domains:**

amino acids 23-42 (type II), 60-80, 97-117, 128-148

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 239**

GTTGATGGCAAACCTTCCTCAAAGGAGGGGCAGAGCCTGCGCAGGGCAGGAGCAGCTGGCCCA  
 CTGGCGGCCCCGCAACACTCCGTCTCACCTCTGGGCCCCTGCATCTAGAGGAGGGCCGTCT  
 GTGAGGCCACTACCCCTCCAGCAACTGGGAGGTGGGACTGTCAGAAGCTGGCCCAGGGTGGT  
 GGTCAGCTGGGTGAGGGACCTACGGCACCTGCTGGACCACCTCGCCTTCTCCATCGAAGCAG  
 GGAAGTGGGAGCCTCGAGCCCTCGGGTGGGAAGCTGACCCCAAGCCACCCTTCACCTGGACAG  
**GATG**AGAGTGTGAGGTGTGCTTCGCCTCCTGGCCCTCATCTTTGCCATAGTCACGACATGGA  
 TGTATTATTCGAAGCTACATGAGCTTCAGCATGAAAACCATCCGTCTGCCACGCTGGCTGGCA  
 GCCTCGCCACCAAGGAGATCCAGGTTAAAAAGTACAAGTGTGGCCTCATCAAGCCCTGCCC  
 AGCCAACTACTTTGCGTTTAAAATCTGCAGTGGGGCCGCCAACGTCGTGGGCCCTACTATGT  
 GCTTTGAAGACCGCATGATCATGAGTCCTGTGAAAAACAATGTGGGCAGAGGCCTAAACATC  
 GCCCTGGTGAATGGAACCACGGGAGCTGTGCTGGGACAGAAGGCATTTGACATGTACTCTGG  
 AGATGTTATGCACCTAGTGAAATTCCTTAAAGAAATTCCGGGGGGTGCAGTGGTGCTGGTGG  
 CCTCCTACGACGATCCAGGGACCAAATGAACGATGAAAGCAGGAAACTCTTCTCTGACTTG  
 GGGAGTTCCTACGCAAAACAACTGGGCTTCCGGGACAGCTGGGTCTTCATAGGAGCCAAAGA  
 CCTCAGGGGTAAAAGCCCCCTTTGAGCAGTTCTTAAAGAACAGCCCAGACACAAACAAATACG  
 AGGGATGGCCAGAGCTGCTGGAGATGGAGGGCTGCATGCCCCGAAGCCATTT**TAG**GGTGGC  
 TGTGGCTCTTCCTCAGCCAGGGGCCTGAAGAAGCTCCTGCCTGACTTAGGAGTCAGAGCCCC  
 GCAGGGGCTGAGGAGGAGGAGCAGGGGGTGCTGCGTGGAAGGTGCTGCAGGTCCTTGACGC  
 TGTGTCGCGCCTCTCCTCCTCGGAAACAGAACCCTCCCACAGCACATCCTACCCGGAAGACC  
 AGCCTCAGAGGGTCCTTCTGGAACCAGCTGTCTGTGGAGAGAATGGGGTGCTTTTCGTCAGGG  
 ACTGCTGACGGCTGGTCCTGAGGAAGGACAACTGCCCAGACTTGAGCCCAATTAAATTTTA  
 TTTTGTGCTGGTTTTGAAAAAAAAAAAAAAAAAAAAA

**FIGURE 240**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59814

<subunit 1 of 1, 224 aa, 1 stop

<MW: 24963, pI: 9.64, NX(S/T): 1

MRVSGVLRLLALIFAIVTTWMFIRSYMFSMKTIRLPRWLAASPTKEIQVKKYKCGLIKPCP  
ANYFAFKICSGAANVVGPTMCFEDRMIMSPVKNNVGRGLNIALVNGTTGAVLGQKAFDMYSG  
DVMHLVKFLKEIPGGALVLVASYDDPGTKMNDESRLKFSDLGSSYAKQLGFRDSWVFIGAKD  
LRGKSPFEQFLKNSPDTNKEYEGWPELLEMEGCMPPKPF

**Important features:****Signal peptide:**

amino acids 1-15

**ATP/GTP-binding site motif A (P-loop).**

amino acids 184-191

**N-glycosylation site.**

amino acids 107-110

**FIGURE 241**

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTTGTCCTGGGGAT  
 CCAGAAACCCATGATACCCTACTGAACACCGAATCCCCTGGAAGCCCACAGAGACAGAGACA  
 GCAAGAGAAGCAGAGATAAATACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCTCTC  
 TCACTCCTCCCTCCCTCTCTCTCTGCTGTCTAGTCCTCTAGTCCTCAAATTTCCAGTCCC  
 CTGCACCCCTTCCTGGGACACTATGTTGTTCTCCGCCCTCCTGCTGGAGGTGATTTGGATCC  
 TGGCTGCAGATGGGGGTCAACACTGGACGTATGAGGGCCACATGGTCAGGACCATTGGCCA  
 GCCTCTTACCCTGAGTGTGGAAACAATGCCCAGTCGCCCATCGATATTCAGACAGACAGTGT  
 GACATTTGACCCTGATTTGCCTGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGC  
 CTTTGGACCTGCACAACAATGGCCACACAGTGCAACTCTCTCTGCCCTCTACCCTGTATCTG  
 GGTGGACTTCCCCGAAAATATGTAGCTGCCAGCTCCACCTGCACTGGGGTCAGAAAGGATC  
 CCCAGGGGGGTGAGAACACCAGATCAACAGTGAAGCCACATTTGCAGAGCTCCACATTGTAC  
 ATTATGACTCTGATTCCTATGACAGCTTGAGTGAGGCTGCTGAGAGGCCTCAGGGCCTGGCT  
 GTCCTGGGCATCCTAATTGAGGTGGGTGAGACTAAGAATATAGCTTATGAACACATTCTGAG  
 TCACTTGCAATGAAGTCAGGCATAAAGATCAGAAGACCTCAGTGCCTCCCTTCAACCTAAGAG  
 AGCTGCTCCCCAAACAGCTGGGGCAGTACTTCCGCTACAATGGCTCGCTCACAACCTCCCCCT  
 TGCTACCAGAGTGTGCTCTGGACAGTTTTTTTATAGAAGGTCCCAGATTTCAATGGAACAGCT  
 GGAAAAGCTTCAGGGGACATTGTTCTCCACAGAAGAGGAGCCCTCTAAGCTTCTGGTACAGA  
 ACTACCGAGCCCTTCAGCCTCTCAATCAGCGCATGGTCTTTGCTTCTTTTCATCCAAGCAGGA  
 TCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGTAGGAATCTTGGTTGGCTGTCTCTG  
 CCTTCTCCTGGCTGTTTATTTTCATTGCTAGAAAGATTCGGAAGAAGAGGCTGGAAAACCGAA  
 AGAGTGTGGTCTTCACCTCAGCACAAGCCACGACTGAGGCATAAATTCCTTCTCAGATACCA  
 TGGATGTGGATGACTTCCCTTCATGCCTATCAGGAAGCCTCTAAAATGGGGTGTAGGATCTG  
 GCCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCCTTCCCCTGGACATCTCTTAGAGAG  
 GAATGGACCCAGGCTGTCATTCCAGGAAGAACTGCAGAGCCTTCAGCCTCTCCAAACATGTA  
 GGAGGAAATGAGGAAATCGCTGTGTTGTTAATGCAGAGANCAAACCTCTGTTTAGTTGCAGGG  
 GAAGTTTGGGATATACCCCAAAGTCCTCTACCCCTCACTTTTATGGCCCTTTCCTAGATA  
 TACTGCGGGATCTCTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGTATATTTTTGATCAATA  
 TATTTGGAAATTAAAGTTTCTGACTTT



**FIGURE 242**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62812

><subunit 1 of 1, 337 aa, 1 stop

><MW: 37668, pI: 6.27, NX(S/T): 1

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWHPASYPECGNNAQSPIDIQTDSVTFDPDLP  
 ALQPHGYDQPGTEPLDLHNNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQ  
 INSEATFAELHIVHYDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRH  
 KDQKTSVPPFNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGT  
 LSTEEEPSKLLVQNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYF  
 IARKIRKKRLENRKSVVFTSAQATTEA

**Important features of the protein:**

**Signal peptide:**

amino acids 1-15

**Transmembrane domain:**

amino acids 291-310

**N-glycosylation site.**

amino acids 213-216

**Eukaryotic-type carbonic anhydrases proteins**

amino acids 197-245, 104-140, 22-69

## FIGURE 243

[illegible]

## **FIGURE 244**

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM  
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE  
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG  
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

**Signal peptide:**

amino acids 1-16

protein = 248 amino acids

**FIGURE 245**

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCG  
GAGCCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCCG  
GCAGCCGGGAGCC**ATG**CGACCCAGGGCCCCGCGCCTCCCCGCAGCGGCTCCGCGGCCTCC  
TGCTGCTCCTGCTGCTGCAGCTGCCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAG  
CAAAAGGCGCAGCTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGG  
GCCAGCAGGAGTGCCTGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTG  
GGATCCCAGGTCGGGATGGATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTTGAG  
GAGTCCTGGACACCCAATAACAAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCT  
TGGGAAAATTGCGGAGTGCTACATTTACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGT  
TCAGTGGCTCACTTCGGCTAAAATGCAGAAATGCATGCTGTCAGCGTTGGTATTTACATTC  
AATGGAGCTGAATGTTCAAGACCTCTTCCCATTTGAAGCTATAATTTATTTGGACCAAGGAAG  
CCCTGAAATGAATTCACAATTAATATTCATCGCACTTCTTCTGTGGAAGGACTTTGTGAAG  
GAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGGTTGGCACTTGTTTCAGATTACCCAAAA  
GGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATCATTATTGAAGAACTACCAAA**TA**  
**A**ATGCTTTAATTTTCATTTGCTACCTCTTTTTTTTATTATGCCTTGGAATGGTTCACTTAAAT  
GACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAGCTAAATATGTTTACAGA  
CCAAAGTGTGATTTACACTGTTTTTAAATCTAGCATTATTCATTTTGCTTCAATCAAAGT  
GGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATTCTCTCAACCTA  
TAATTTGGAATATTGTTGTGGTCTTTTGTCTTTTCTCTTAGTATAGCATTTTTTAAAAAATA  
TAAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGTAAAT  
AAAAATTATTTCCAACA

### FIGURE 246

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76393
```

><subunit 1 of 1, 243 aa, 1 stop

><MW: 26266, pI: 8.43, NX(S/T): 1

MRPQGPAASPQRLRGLLLLLLLLQLPAPSSASEIPK GKQKAQLRQREVVDLYNGMCLQGPAGV  
PGRDGSPGANVIPGTPGIPGRDGFKEGEGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIA  
ECTFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAI IYLDQGSPEMN  
STINIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIIIEELPK

**Signal peptide:**

amino acids 1-30

**Transmembrane domain:**

amino acids 195-217